

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2005, 16:08:36 ; Search time 4402 Seconds
(without alignments)
10570.999 Million cell updates/sec

Title: US-09-830-810A-1

Perfect score: 1277

Sequence: 1 aagcggcgcaggcgcgga.....acaaaaaAAAAAAAAA 1277

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gsl1:*

9: gb_gsl2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	643	50.4	675	2	BB641267
2	642.4	50.3	666	2	BB264222
3	603.4	47.3	845	7	CO800921
C	591.4	46.3	615	4	EG071693
5	582.4	45.6	595	4	EG084538
6	571.4	44.7	731	7	CO809618
C	570.4	44.7	594	2	BF471866
8	522.8	40.9	722	7	CO815317
9	502.4	39.3	546	7	CO797395
10	497.2	38.9	525	2	BB704019
11	480.4	37.6	687	7	CO804482
C	473.8	37.1	499	1	AI854700
13	448.4	35.1	491	2	BB703259
14	413.8	32.4	491	2	BB704648
C	405.8	31.8	436	1	AU044294
16	398.2	31.2	425	2	BB705931
C	396.4	31.0	590	1	AU023153
18	393	30.8	450	2	BB946858
19	392.2	30.7	419	2	BB704449
20	367.4	28.8	521	2	BB703869
21	367	28.7	427	2	BB706957
22	364.2	28.5	434	2	BB699732
23	347.2	27.2	391	2	BB704146
24	333.8	26.1	400	2	BB700620

25	319	25.0	905	8	BZ114495
26	316.2	24.8	441	2	BB704848
C	282.2	22.1	301	5	BX637052
28	284.6	20.7	315	2	BB248342
29	227.4	17.8	642	6	BY179704
30	224.8	17.6	654	8	AZ225053
31	211.6	16.6	251	1	AV359580
32	208.6	16.3	251	1	AV358045
C	207.4	16.2	819	5	BP740232
34	203.2	15.9	249	1	AV357766
35	203	15.9	290	2	BB063113
C	201.8	15.8	539	1	AL904756
C	200.8	15.7	909	5	EX755136
38	200	15.7	286	2	BB060381
39	199.8	15.6	246	2	BB060134
C	198	15.5	539	1	AL904892
C	197.8	15.5	642	4	EM861479
C	197	15.4	643	4	BU098465
C	196.8	15.4	526	1	AL904839
C	196.8	15.4	534	1	AL904820
C	196.8	15.4	559	1	AL904774

ALIGNMENTS

RESULT 1	BB641267	675 bp	linear	EST 26-OCT-2001
LOCUS	BB641267	musculus	CDNA clone A830014H23 5', mRNA sequence.	
DEFINITION	BB641267	musculus	CDNA clone A830014H23 5', mRNA sequence.	
ACCESSION	BB641267	GI:16476392		
VERSION	BB641267.1	GI:16476392		
KEYWORDS	EST.			
SOURCE	Mus musculus			
ORGANISM	Mus musculus			
REFERENCE	1 (bases 1 to 675)			
AUTHORS	Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.			
TITLE	RIKEN Mouse ESTs (Arakawa, T., et al. 2001)			
JOURNAL	Unpublished (2001)			
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/ Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)			

FEATURES
source

Db	62	CATCTTATTCGCGACGGCACAAAGCCGGGATGGCTGGAGGTTTCGAGCCGACGGGCGTGC	121
Qy	121	CGACCGCGCGCCCCCTCTCTCTCCCGGGTACAGACAGCTCATGGCCGCGAGGTACGTC	180
Db	122	CGACCGCGCGCCCCCTCTCTCTCCCGGGTACAGACAGCTCATGGCCGCGAGGTACGTC	181
Qy	181	GACAGCCACGAGGGGACAGCTCATGCCCTGCTGCGCGATGGGTCCCGGTCGGTC	240
Db	182	GACAGCCACGAGGGGACAGCTCATGCCCTGCTGCGCGATGGGTCCCGGTCGGTC	241
Qy	241	AGCAGCCGTGACGCTGCGGTGAGGTAAACCCGCGCGCGACGCTCGGTGCAGTGTTC	300
Db	242	AGCAGCCGTGACGCTGCGGTGAGGTAAACCCGCGCGCGACGCTCGGTGCAGTGTTC	301
Qy	301	CTGCGGGCGCGACGCTCGAGCCTTCGAGGTTCCGAGCAGCCCGACGCCGATCGGCT	360
Db	302	CTGCGGGCGCGACGCTCGAGCCTTCGAGGTTCCGAGCAGCCCGACGCCGATCGGCT	361
Qy	361	TCCTGTCAACCCCGTGGCCACGCCCGCGCGCGGAGATCCCGCGATCCTGGCAGACGTA	420
Db	362	TCCTGTCAACCCCGTGGCCACGCCCGCGCGCGGAGATCCCGCGATCCTGGCAGACGTA	421
Qy	421	GCCCGGTTCTGTCGTCGAGACCTTCTGTGGCCTCTCTCTCACTGAGGTTGCGGAGGC	480
Db	422	GCCCGGTTCTGTCGTCGAGACCTTCTGTGGCCTCTCTCTCACTGAGGTTGCGGAGGC	481
Qy	481	AGGCAGACACCCACGAAGGAGAGGGGAGCCCGGATCCTCGGGACCCGGGACCGGAG	540
Db	482	AGGCAGACACCCACGAAGGAGAGGGGAGCCCGGATCCTCGGGACCCGGGACCGGAG	541
Qy	541	CCGAGAGAGTGGCCCGAGGAAAGCGGTCCCGACGCCGGAAGCAGAGGAGCGCATGTT	600
Db	542	CCGAGAGAGTGGCCCGAGGAAAGCGGTCCCGACGCCGGAAGCAGAGGAGCGCATGTT	600
Qy	601	CAGGCTGCAGGGCAGGCGGGTGGGAGCA	629
Db	601	CAGGCTGCAGGGC-GGCCGGGTGGGAGCA	628
RESULT 4			
BG071693/c			
LOCUS			
DEFINITION			
H3102B03-3 NTA Mouse 15K cDNA Clone Set Mus musculus cDNA clone			
H3102B03 3', mRNA sequence.			
ACCESSION			
VERSION			
BG071693			
BG071693.2 GI:40072037			
EST.			
KEYWORDS			
Mus. musculus (house mouse)			
EST.			
Mus. musculus (house mouse)			
EST.			
Mus. musculus (house mouse)			
EST.			
Mus. musculus (house mouse)			
EST.			
Mus. musculus (house mouse)			
EST.			
Mus. musculus (house mouse)			
EST.			
Mus. musculus (house mouse)			
EST.			
Mus. musculus (house mouse)			
EST.			
Mus. musculus (house mouse)			
EST.			
Mus. musculus (house mouse)			
EST.			
Mus. musculus (house mouse)			
EST.			
Mus. musculus (house mouse)			
EST.			
Mus. musculus (house mouse)			
EST.			
Mus. musculus (house mouse)			
EST.			
Mus. musculus (house mouse)			
EST.			
Mus. musculus (house mouse)			
EST.			
Mus. musculus (house mouse)			
EST.			
Mus. musculus (house mouse)			
EST.			
Mus. musculus (house mouse)			
EST.			
Mus. musculus (house mouse)			
EST.			
Mus. musculus (house mouse)			
EST.			
Mus. musculus (house mouse)			
EST.			
Mus. musculus (house mouse)			
EST.			

RESULT 4	615 bp	mRNA	linear	EST 18-DEC-2003
BG071693/c				
LOCUS				
DEFINITION	H3102B03-3 NTA Mouse 15K cDNA Clone Set Mus musculus cDNA clone			
ACCESSION	H3102B03 3', mRNA sequence.			
VERSION	BG071693			
KEYWORDS	BG071693.2 GI:40072037			
SOURCE	EST.			
ORGANISM	Mus musculus (house mouse)			
REFERENCE	Mus musculus			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 615)			
TITLE	Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Granhovac,M.J., Pantano,S., Sano,Y., Piao,Y., Ngataraja,R., Doi,H., Wood,W.H. III, Becker,K.G. and Ko,M.S.H. Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)			
JOURNAL	20381348.			
MEDLINE	10922068			
PUBLISHED	On Jan 26, 2001 this sequence version replaced gi:12554262.			
COMMENT	Other_ESTs: H3102B03-5 Contact: George J. Kargul Laboratory of Genetics National Institute on Aging/National Institutes of Health 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA Email: cdna@lgsun.grc.nia.nih.gov This clone set has been freely distributed to the community. Please visit http://lgsun.grc.nia.nih.gov/cDNA/15k.html for details. Plate: H3102 row: B column: 03 Seq primer: -21M13 Forward High quality sequence stop: 615 POLYA=Yes.			

Location/Qualifiers
1. .615

FEATURES	source
1. .615	
Location/Qualifiers	
/organism="Mus musculus"	
/mol_type="mRNA"	
/strain="C57BL/6J"	
/db_xref="niaEST:H3102B03-3"	
/db_xref="taxon:10090"	
/clone="H3102B03"	
/sex="Clones arrayed from a variety of cDNA libraries"	
/dev_stage="Clones arrayed from a variety of cDNA libraries"	
lab_host="DH10B"	
/clone_lib="NIA Mouse 15K cDNA Clone Set"	
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This clone is among a rearrayed set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on chr. X chromosome, 1998, Hum Mol Genet 7: 1967-1978."	

ORIGIN

Query Match	46.3%	Score 591.4	DB 4	Length 615
Best Local Similarity	99.5%	Pred. No. 1.7e-133		
Matches 614; Conservative	0	Mismatches 1	Indels 2	Gaps 2
643	Y	GAGGACCGGAAACAGTGTGGCGGCATGCAGTCTGAGCCTGGGACGAGGAGCATGTCT	702	
615	b	GAGGACCGGAACAGTGTGGCGCATGCAGTCTGAGCCTGGGACGAGGAGCATGTCT	556	
703	Y	GCCGACAGATGGCTCAGGACCCCGGTGAATTGGATGCCCTCGAACACAGGCTCCCG	762	
555	b	GCCGACAGATGGCTCAGGACCCCGGTGAATTGGATGCCCTCGAACACAGGCTCCCG	496	
763	Y	CAAGACGCGACGAGCAAGAGGCGCTGGTTTCCAGTCTTTAGAGCAAGAATGACGCG	822	
495	b	CAAGCACGGACGAGCAAGAGGCGCTGGTTTCCAGTCTTTAGAGCAAGAATGACGCG	436	
823	Y	TACTATCACTGCAAGGACTGCAAAATCCGGTGGGAGAGCGCCTATGTGTGTGTGCA	882	
435	b	TACTATCACTGCAAGGACTGCAAAATCCGGTGGGAGAGCGCCTATGTGTGTGTGCA	376	
883	Y	GGCACCAAGTAAAGTGTACTTCAAAACAGTTCTGCCGAGTGTGAGAAATCTTACAACC	942	
375	b	GGCACCAAGTAAAGTG-TACTTCAAAACAGTTCTGCCGAGTGTGAGAAATCTTACAACC	317	
943	Y	TTACAGAGTGGAGACATCACTGTCAAGTTGTATAAGAACTAGATGTGCCTGCCAGT	1002	
316	b	TTACAGAGTGGAGACATCACTGTCAAGTTGTATAAGAACTAGATGTGCCTGCCAGT	257	
1003	Y	CAGATTTGCGCAGTGGACCTTAAACGCCCCCATCGGCAAGACTTGTGTGGGAGATGCAA	1062	
256	b	CAGATTTGCGCAGTGGACCTTAAACGCCCCCATCGGCAAGACTTGTGTGGGAGATGCAA	197	
1063	Y	GGACAAACGCTCTCTGCGACAGCACCTTACGCTTCAATACATCATTTAGTGAGAGTC	1122	
196	b	GGACAAACGCTCTCTCTGCGACAGCACCTTACGCTTCAATACATCATTTAGTGAGAGTC	137	
1123	Y	GAACAAGCTTTCGCTAGATGGGCTTAATGGAATGCAAGTGAAGCTTTCTCCCTCTTCA	1182	
136	b	GAACAAGCTTTCGCTAGATGGGCTTAATGGAATGCAAGTGAAGCTTTCTCCCTCTTCA	77	

Qy 1183 CCTCTTCCCTTTCCAAATTTCTTCATGACAGACAGTGTACTTGGATATAAAGCCGTGTAA 1242
|||||
Db 76 CCTCTTCCCTTTCCAAATTTCTTCATGACAGACAGTGTACTTGGATATAAAGCCGTGTAA 18

Qy 1243 TAAAGGTATTGCAAAAC 1259
|||||
Db 17 TAAAGGTATTGCAAAAC 1

RESULT 5
LOCUS BG084538
DEFINITION H3102B03-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone EST 18-DEC-2003
H3102B03 5', mRNA sequence.
ACCESSION BG084538
VERSION BG084538.2 GI:40072038
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 595)
Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X.,
Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H.,
Wood,W.H. III, Becker,K.G. and Ko,M.S.H.
Genome-wide expression profiling of mid-gestation placenta and
embryo using a 15,000 mouse developmental cDNA microarray
Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
20381348
10922068
On Jan 26, 2001 this sequence version replaced gi:12567102.
Other ESTs: H3102B03-3
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit <http://lgsun.grc.nia.nih.gov/cDNA/15k.html> for details.
Plate: H3102 row: B column: 03
Seq primer: -21M13 Reverse
High quality sequence stop: 595
POLYA=No.

FEATURES
source
1..595
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:H3102B03-5"
/db_xref="taxon:10090"
/clone="H3102B03"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stages="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/note="Vector: pSPORT1; Site1: SalI; Site2: NotI; This clone is among a rearrayed set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U.S.A. 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from

ORIGIN

Query Match 45.6%; Score 582.4; DB 4; Length 595;
Best Local Similarity 99.7%; Pred. No. 2.6e-131;
Matches 594; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 511 CCGCATCTCCGGGACCCGGACCGGACCGGAGAGAGGTGGCGCGGAGAAAGCGGTC 570
|||||
Db 1 CCGCATCTCCGGGACCCGGGACCCGGGACCGGAGAGAGGTGGCGCGGAGAAAGCGGTC 60
|||||
Qy 571 CCCAGCGCGCAAGCAGGAGGGCGATGTTTCAGGCTGCAGGCGAGGCGCGGTGGAGCAG 630
|||||
Db 61 CCCAGCGCGCAAGCAGGAGGGCGATGTTTCAGGCTGCAGGCGAGGCGCGGTGGAGCAG 120
|||||
Qy 631 CAGCCACACCGGAGGACCGGAACAGTGTGGCGCGATGTCAGTCTGAGCCTGGGAGCGAG 690
|||||
Db 121 CAGCCACACCGGAGGACCGGAACAGTGTGGCGCGATGTCAGTCTGAGCCTGGGAGCGAG 180
|||||
Qy 691 GAGCCATGTCCTCCGCGAGAGATGGCTCAGGACCCCGGTGATTCGGATGCCCTCGAGAC 750
|||||
Db 181 GAGCCATGTCCTCCGCGAGAGATGGCTCAGGACCCCGGTGATTCGGATGCCCTCGAGAC 240
|||||
Qy 751 CAGGCTCCCGCAAGCAGGAGCAGGACAAAGAGAGCGCCTGCGTTTCCAGTCTCTTAGAG 810
|||||
Db 241 CAGGCTCCCGCAAGCAGGAGCAGGACAAAGAGAGCGCCTGCGTTTCCAGTCTCTTAGAG 300
|||||
Qy 811 CAGAAGTACGGCTACTATCTACTGCAAGACTGCAAAATCCGGTGGGAGAGCGCCTATGTG 870
|||||
Db 301 CAGAAGTACGGCTACTATCTACTGCAAGACTGCAAAATCCGGTGGGAGAGCGCCTATGTG 360
|||||
Qy 871 TGGTGTGTCAGGCGCACCACTAAGTGTACTTCAAACAGTTCTGCCGAGTGTGTGAGAA 930
|||||
Db 361 TGGTGTGTCAGGCGCACCACTAAGTGTACTTCAAACAGTTCTGCCGAGTGTGTGAGAA 419
|||||
Qy 931 ATCTCTACAACCTTTACAGAGTGGAGGACATCACCTGTCAAAAGTTGTAAAGAACTAGATG 990
|||||
Db 420 ATCTCTACAACCTTTACAGAGTGGAGGACATCACCTGTCAAAAGTTGTAAAGAACTAGATG 479
|||||
Qy 991 TGCTTGCCAGTCTAGATTTTCCGACAGTGGACCTTAAAGCGCCCATGGCAAGACTTTGTG 1050
|||||
Db 480 TGCTTGCCAGTCTAGATTTTCCGACAGTGGACCTTAAAGCGCCCATGGCAAGACTTTGTG 539
|||||
Qy 1051 TGGGAGATGCAAGGACAAACGCCCTGCTCTGCGAGCAGCACCTTCAGCTTCAAAATACA 1106
|||||
Db 540 TGGGAGATGCAAGGACAAACGCCCTGCTCTGCGAGCAGCACCTTCAGCTTCAAAATACA 595
|||||

RESULT 6

LOCUS CO809618
DEFINITION AGENCOURT 30247201 NIH MGC 256 Mus musculus cDNA clone
IMAGE:30936113 5', mRNA sequence.
ACCESSION CO809618
VERSION CO809618.1 GI:51027468
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 731)
REFERENCE
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgaps-remail.nih.gov
Tissue Procurement: Dr. Kathleen Horner, Stanford University
cDNA Library Preparation: Express Genomics


```
Db 594 GCGCAGGAAAGCGGTCCCGCCAGCGCGGAAAGCGAGGCGCGATGTTCAAGGCTGCAGGGCA 535
Qy 615 GCGCGGTGGGAGCAGCAGCACCACCGAGGACCGGACAGTGTGGCGGCGATGCAGTTC 674
Db 534 GCGCGGTGGGAGCAGCAGCACCACCGAGGACCGGACAGTGTGGCGGCGATGCAGTTC 475
Qy 675 TGAGCCTGGGAGCAGGAGCCATGTCCTCCGCGAGAGATGGCTCAGGACCCCGGTGATTC 734
Db 474 TGAGCCTGGGAGCAGGAGCCATGTCCTCCGCGAGAGATGGCTCAGGACCCCGGTGATTC 415
Qy 735 GGATGCCCTCGAGACGAGGCTCCCGCAAGACCGAGAGGACGAGAGCGGCTCGG 794
Db 414 GGATGCCCTCGAGACGAGGCTCCCGCAAGACCGAGAGGACGAGAGCGGCTCGG 355
Qy 795 TTTCCAGTCTTAGAGCAGAGTAGTCCGCTACTATCATCTGCAAGGACTCAAAATCCGGTG 854
Db 354 TTTCCAGTCTTAGAGCAGAGTAGTCCGCTACTATCATCTGCAAGGACTCAAAATCCGGTG 295
Qy 855 GGAGAGCGCCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 914
Db 294 GGAGAGCGCCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 236
Qy 915 GCGAGTGTGTGAGAAATCTTACAAACCTTACAGAGTGGAGACATCACTGTCAAAGTT 974
Db 235 GCGAGTGTGTGAGAAATCTTACAAACCTTACAGAGTGGAGACATCACTGTCAAAGTT 176
Qy 975 GTAAAGAACTAGATGTCCTGCCAGTCCAGTTCAGATTCGCGACGAGCCTAAACGCCCC 1034
Db 175 GTAAAGAACTAGATGTCCTGCCAGTTCAGATTCGCGACGAGCCTAAACGCCCC 116
Qy 1035 ATCGCAAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1094
Db 115 ATCGCAAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 56
Qy 1095 GCTTCAAAATACATCATTTAGTAGAGTGCAGAAACGTTTCTGTGTGTGTGTGTGTGT 1150
Db 55 GCTTCAAAATACATCATTTAGTAGAGTGCAGAAAC-TTTCTGTGTGTGTGTGTGTGTGT 1
```

```
RESULT 8
COB15317 722 bp mRNA linear EST 06-AUG-2004
LOCUS AGNCOURT 30256635 NIH MGC 256 Mus musculus cDNA clone
DEFINITION IMAGE:30939000 5', mRNA sequence.
ACCESSION COB15317
VERSION COB15317.1 GI:51033964
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
```

```
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 722)
JOURNAL NIH-MGC http://mgi.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Kathleen Horner, Stanford University
cDNA Library Preparation: Express Genomics
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM1198 row: n column: 01
High quality sequence stop: 495.
Location/Qualifiers
1..722
/organism="Mus musculus"
```

```
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30939000"
/lab_host="DH10B_TonA"
/clone_lib="NIH_MGC_256"
/note="Organ: oocyte; Vector: pExpress-1; Site: 1; EcoRV;
Site 2: NotI; cDNA was primed using oligo-dT primer:
5'-pGACTAGTTCATGATCGAGCGGCCCT(T)25-3' and cloned into
the EcoRV/NotI sites of pExpress-1. Size-selection >0.5 kb
resulted in an average insert size of 1.2 kb. This is a
primary library (normalized insert size) is NIH_MGC_257)
and was constructed by Express Genomics (Frederick, MD).
Note: this is a NIH_MGC library"
```

ORIGIN

```
Query Match 40.9%; Score 522.8; DB 7; Length 722;
Best Local Similarity 91.8%; Pred. No. le-116;
Matches 561; Conservative 0; Mismatches 47; Indels 3; Gaps 1;

Qy 7 GCGCAGGCGCGGAGCAGCACCCTGTTCCCGCGAGCAGCTTCCACCCCTGCCGATCCT 66
Db 4 GCGCAGGCGCGGAGCAGCACCCTGTTCCCGCGAGCAGCTTCCACCCCTGCCGATCCT 63

Qy 67 TATCCGAGGCGCACAAAGCCGGGATGGCTGGAGGTTTCGAGCCAGGGCTGCCAGCC 126
Db 64 TATCCGAGGCGCACAAAGCCGGGATGGCTGGAGGTTTCGAGCCAGGGCTGCCAGCC 123

Qy 127 GCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 186
Db 124 GCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 183

Qy 187 CACCAGCGGCGCACAGCTCATGGCCCTGCTGTGCGGAGTGGTCCCGCTCGAGCAGC 246
Db 184 CACCAGCGGCGCACAGCTCATGGCCCTGCTGTGCGGAGTGGTCCCGCTCGAGCAGC 243

Qy 247 CGTACGCTGCGGTGCAGAGTGAAACCGCGCCCGCGAGCGCTCGGTGCAAGTTCAC 306
Db 244 CGTACGCTGCGGTGCAGAGTGAAACCGCGCCCGCGAGCGCTCGGTGCAAGTTCAC 303

Qy 307 GCGCGCAGCTGCAGCCTGCGAGGTCGCGAGCCAGCCCGCGAGCCCGATCGGTTCTGT 366
Db 304 GCGCGCAGCTGCAGCCTGCGAGGTCGCGAGCCAGCCCGCGAGCCCGATCGGTTCTGT 363

Qy 367 CAACCCCGTGCACACGCGCGCGCGGAGATCCCGCGATCCTGCGAGACCGTAGCCCG 426
Db 364 CAACCCCGTGCACACGCGCGCGCGGAGATCCCGCGATCCTGCGAGACCGTAGCCCG 423

Qy 427 TTCTCGTCCGTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 486
Db 424 TTCTCGTCCGTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 483

Qy 487 ACACCCAGAGGAGGAGGAGCGCGCGATCCTCGGGACCCCGGAGACCGGAGCCGAGA 546
Db 484 ACACCCAGAGGAGGAGGAGGAGCGCGATCCTCGGGAGACCGGAGACCGGAGCCGAGA 540

Qy 547 GAGGTGCGCGGAGGAGGAGGAGCGCGCGATCCTCGGGAGGAGGAGGAGGAGGAGGAG 606
Db 541 GAGGTGCGCGGAGGAGGAGGAGGAGCGCGCGATCCTCGGGAGGAGGAGGAGGAGGAG 600

Qy 607 GCAGGGCAGGC 617
Db 601 GCGCGGTGGAC 611
```

RESULT 9

```
CO797395
LOCUS AGNCOURT 30246265 NIH MGC 257 Mus musculus cDNA clone
DEFINITION IMAGE:30935142 5', mRNA sequence.
ACCESSION CO797395
VERSION CO797395.1 GI:50985575
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
```

FEATURES

source

CO797395 546 bp mRNA linear EST 05-AUG-2004
AGNCOURT 30246265 NIH MGC 257 Mus musculus cDNA clone
IMAGE:30935142 5', mRNA sequence.

ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 546)
AUTHORS	NIH-MGC http:// mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics / NIH National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: cgapbe-remail.nih.gov Tissue Procurement: Dr. Kathleen Horner, Stanford University CDNA Library Preparation: Express Genomics DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: NDA01188 row: m column: 07 High quality sequence stop: 519. Location/Qualifiers 1..546 /organism="Mus musculus" /mol_type="mRNA" /db_xref="taxon:10090" /clones="IMAGE:30935142" /lab_host="DH10B Tona" /clone_lib="NIH_MGC_257" Note: "Organ: oocyte; Vector: pExpress-1; Site: 1: EcoRV; Site 2: NotI; CDNA was primed using oligo-dr primer: 5'-pGACTAGTTCGGACGTCCGCC(T)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. Size-selection >0.5 kb resulted in an average insert size of 1.0kb. This is a normalized library (primary library is NIH_MGC_256) and was constructed by Express Genomics (Frederick, MD). Note: this is a NIH_MGC library"
FEATURES	source
ORIGIN	Query Match 39.3%; Score 502.4; DB 7; Length 546; Best Local Similarity 98.5%; Pred. No. 9.2e-112; Matches 525; Conservative 0; Mismatches 6; Indels 2; Gaps 2;
QY	745 CGAGACCAGCGCTCCCGCAAAGCACGGAGCAGGACAAGAGCGGCTTGCTTCCAGTTTC 804
Db	1 CGAGACCAGCGCTCCCGCAAAGCACGGAGCAGGACAAGAGCGGCTTGCTTCCAGTTTC 60
QY	805 TTAGAGCAGAAGTAGCGGTACTATCACTGCAAGGACTGCAAAATCGGTGGAGAGCGCC 864
Db	61 TTAGAGCAGAAGTAGCGGTACTATCACTGCAAGGACTGCAAAATCGGTGGAGAGCGCC 120
QY	865 TATGTGTGTGTGTGCAGGGCCACCAGTAAGGTGTACTTCAAACAGTTCTGCCGAGTGTG 924
Db	121 TATGTGTGTGTGTGCAGGGCCACCAGTAAGGTG-TACTTCAAACAGTTCTGCCGAGTGTG 179
QY	925 TGAGAAATCCTAACACCCATTACAGAGTGAGAGCATCACTGTCAAAGTTGTTAAAAAGAC 984
Db	180 TGAGAAATCCTAACACCCATTACAGAGTGAGAGCATCACTGTCAAAGTTGTTAAAAAGAC 239
QY	985 TAGATGTGCTGCCAGTCAGATTTCCGCACGTGGACCTTAACAGCCCCCATCGCAAGA 1044
Db	240 TAGATGTGCTGCCAGTCAGATTTCCGCACGTGGACCTTAACAGCCCCCATCGCAAGA 299
QY	1045 CTTGTGTGGAGATGCAAGGACAAACCGCTGTCCTCGCAGCAGCACCTTCAGCTTCAAATA 1104
Db	300 CTTGTGTGGAGATGCAAGGACAAACCGCTGTCCTCGCAGCAGCACCTTCAGCTTCAAATA 359
QY	1105 CATCATTTAGTAGAGTCGAAAACGTTTCTGCTAGATGGGGCTTAATGGAATGACAAGTG 1164
Db	360 CATCATTTAGTAGAGTCGAAAACGTTTCTGCTAGATGGGGCTTAATGGAATGACAAGTG 419
QY	1165 AGCTTTCTCCCTCTTCACCTCTTCCCCTTTCCAAATTTCTTCATGACAGACAGTGTACTTT 1224

<hr/>					
ACCESSION	AU044294				
VERSION	AU044294.1	GI:3979844			
KEYWORDS	EST.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus-				
	1 (bases 1 to 436)				
	Ko, M.S.H., Kitchen, J.R., Wang, X., Threat, T.A., Sun, T.,				
	DePalma, G.E., Liang, Y., Kargul, G.J., Sharara, R., Lim, M.K. and				
	Doi, H.				
TITLE	Systematic analyses of genes expressed in 16-cell mouse embryo (The ERATO/Doi Project at Wayne State University)				
JOURNAL	Unpublished (1998)				
COMMENT	Contact: Hirofumi Doi				
	Doi Biosymmetry Project, ERATO				
	Japan Science and Technology Corporation (JST)				
	WBG Marine East 12F, 2-6 Nakase, Mihama-ku, Chiba 261-71, Japan				
	Email: hd@bioa.jst.go.jp.				
FEATURES	Location/Qualifiers				
source	1..436				
	/organism="Mus musculus"				
	/mol_type="mRNA"				
	/strain="C57BL/6J"				
	/db_xref="taxon:10090"				
	/clone="J0917G09"				
	/dev_stage="sixteen-cell-embryo"				
	/clone_lib="Mouse sixteen-cell-embryo cDNA"				
ORIGIN					
Query Match	31.8%;	Score 405.8;	DB 1;	Length 436;	
Best Local Similarity	97.0%;	Pred No. 3.6e-88;			
Matches 424;	Conservative	0;	Mismatches 12;	Indels 1;	Gaps 1;
QY	823	TACTATCACTGCAGAGGACTGCAAAATCCGGTGGAGAGCGCCTATGTGTGTGTGTGCAG	882		
Db	436	TATTATCACTGCAGAGGACTGCAAAATCCGGTGGAGAGCGCCTATGTGTGTGTGTGCAG	377		
QY	883	GGCACCCAGTAGGTGTACTTCAAACAGTTCTGCCGAGTGTTGCAGAAATCCTCAACACC	942		
Db	376	GGCACCCAGTAGGTGTACTTCAAACAGTTCTGCCGAGTGTTGCAGAAATCCTCAACACC	317		
QY	943	TTACAGAGTGGAGGACATCATCTGTCAAAGTTGTAAAGAACAATACTAGATGTGCTGCCAGT	1002		
Db	316	TTACAGAGTGGAGGACATCATCTGTCAAAGTTGTAAAGAACAATACTAGATGTGCTGCCAGT	257		
QY	1003	CAGATTCGCCACGTGGACCTTAACCGCCCCCATCGCAAGACTTGTCTGGGAGATGCAA	1061		
Db	256	CAGACTTCGCCACGTGGACCTTAACCGCCCCCATCGCAAGACTTGTCTGGGAGATGCAA	197		
QY	1063	GGACAAACCGCTGTCTCTCGCAGACAGCACCTTCAGCTTCAAATACATCATTTAGTGAGATC	1122		
Db	196	GGACAAACCGCTGTCTCTCGCAGACAGCACCTTCAGCTTCAAATACATCATTTAGTGAGATC	137		
QY	1123	GAAAAAGTTTTCTGTAGATGGGGCTAATNGGAATGCAAGATGAGCTTTCTCCCTCTTCA	1181		
Db	136	GAAAAAGTTTTCTGTGTAGATGGGGCTAATNGGAATGCAAGATGAGCTTTCTCCCTCTTCA	77		
QY	1183	CTCTTCTCCCTTCTCCAAATTTCTTCATGACAGACAGTGTTACTTTGGATATAAAGCCCTGTGAA	1241		
Db	76	CTCTTCTCCCTTCTCCAAATTTCTTCATGACAGACAGTG-TACTTTGGATATAAAGCCCTGTGAA	18		
QY	1243	TAAAAAGGTATTGCAAC	1259		
Db	17	TAAAAAGGTATTGCAAC	1		

Search completed: January 5, 2005, 19:20:23
Job time : 4411 secs

```

/strain="C57BL/6J"
/db_xref="taxon:10090"
/cloned="742046L07"
/sex="female"
/tissue_type="in vitro fertilized eggs"
/dev_stage="egg"
/lab_host="DH10B"
/cloned_lib="RIKEN full-length enriched, in vitro
fertilized eggs"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGAGGATCCAGAGAGCTCTTTTTTTTTTTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5',
GAGAGAGAGATCTCGAGTAAATTAATTAATTAATTCCTCCCTCCCTCC 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"

```

ORIGIN

Query Match	32.4%;	Score 413.8;	DB 2;	Length 491;
Best Local Similarity	93.0%;	Pred. No. 4.1e-90;		
Matches 455;	Conservative 0;	Mismatches 32;	Indels 2;	Gaps 2;

```

QY      768  CACGGAGCAGCAAGAGCGCTCGGTTTCCAGTTCTTAGACAGCAAGTAGCGGTACTCA 827
DB      1    CACGGAGCAGCAGAGAGCTCTGGTTTCCAGTTCTTAGACAGCAAGTAGCTTACTA 60

QY      828  TCACTGCAAGGACTGCAAAATCCGGTGGGAGAGCGGCTATGTGTGTGTGTCAGGGCAC 887
DB      61  TCA-AGCAAGGACTGCAAAATCCGCTGGGAAAAACGGCTATGTCTGGAGTGTGAGGGCAC 119

QY      888  CAGTAGGTCCTACTTCAACAGCTCTGCGAGTGTGTGAGAAATCTCAACACCTTACA 947
DB      120  CAGTAAAGATGACTTCAACAGTTCTGGCCGAGTGTGTGAGAAATCTCAACACCTTACA 179

QY      948  GAGTGGAGGACATCACCTGTCAAAGTTGTAAAAAGAACTAGATGTGCTGCCACGTGAGAT 1007
DB      180  GAATGGAGGACATCACCTCTCAAATTTGTAAGAAACTAAATGTGCTTCCAGTCAGAC 239

QY      1008  TTGCCACAGTGACCTTAACGCCGCCCATCGGCAAGACTTGTGTGGGAGATGCAAGGACA 1067
DB      240  TTGCCACAGTGACCTTAACGCCGCCCATCGGCAAGACTTGTGTGGGAGATGCAAGGACA 239

QY      1068  AACGCTTCTCTCGCAGCAGACACCTTCAGCTTCAAATACATCATTTAGTGAGAGTCGAAAA 1127
DB      300  AACGCTTCTCTCGCAGCAGCACCCTTCAGCTTCAAATACATCATTTAGTGAGAGTCGAAAA 359

QY      1128  CGTTTCTGTAGATGGGGCTTAATCGAATGGAACAGTCACTTTCTCCCTCTTCACTCT 1187
DB      360  CGTTTCTGTAGATGGGGCTTAATCGAATGGAACAGTCACTTTCTCCCTCTTCACTCT 419

QY      1188  TCCCTTTTCCAAATCTTTCATCACAGACAGTGTACTCTTGGATATAAAGCCTGTGTAATAAA 1247
DB      420  TCCCTTTTCCAAATCTTTCATCACAGACAGTGT-TACTTTGGATATAAAGCCTGTGTAATAAA 478

QY      1248  GGTATTGCA 1256
DB      479  GGTATTGTA 487

```

RESULT 15	linear	EST 09-DEC-1998
AU044294/c	436 bp	
LOCUS	mouse sixteen-cell-embryo	
CDNA	Mus musculus	
DEFINITION	Mouse sixteen-cell-embryo cDNA	
	JO917G09 3', mRNA sequence.	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2005, 14:55:31 ; Search time 5526 Seconds
(without alignments)
10928.148 Million cell updates/sec

Title: US-09-830-810A-1

Perfect score: 1277

Sequence: 1 aagcgcgagcgagcgaggga.....acaaaaaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1245.4	97.5	1260	10 AY191415	Mus musculus
2	962.2	75.3	1280	10 AY283175	Rattus no
3	799.4	62.6	3987	10 AY193889	Mus musculus
4	799.4	62.6	144954	2 AC122733	Mus musculus
5	799.4	62.6	200535	10 AL671880	Mouse DNA
6	689.8	54.0	193274	2 AC107686	Mus musculus
7	628	49.2	235390	2 AC125593	Rattus no
8	628	49.2	237695	2 AC125993	Rattus no
9	628	49.2	262139	2 AC127083	Rattus no
10	578.8	45.3	220894	2 AC108848	Mus musculus
11	352	27.6	1275	9 AY191416	Homo sapi
12	239.2	18.7	366	6 CQ716345	Sequence
13	205.8	16.1	1052	5 AY283176	Xenopus 1
14	194.8	15.3	963	5 AY283177	Takifugu
15	194.6	15.2	1084	5 AY283178	Danio rer
16	147.6	11.6	138890	9 AC007970	Homo sapi
17	147.6	11.6	151961	2 AC037443	Homo sapi
18	147.6	11.6	204268	2 AC092163	Homo sapi
19	134.6	10.5	4138	9 AY193890	Homo sapi

20	134.6	10.5	187578	9	AC096952	Homo sapi
21	97.6	7.6	945	5	BX931269	Gallus ga
22	90.6	7.1	187278	2	BX571888	Danio rer
23	90.6	7.1	249791	5	BX537133	Zebrafish
24	73.2	5.7	694	9	BSA335296	
25	72.2	5.7	125020	9	AF429315	Homo sapi
26	64.8	5.1	144000	9	AF429315	Homo sapi
27	63.2	4.9	125020	9	AF429315	Homo sapi
28	62.6	4.9	1393	11	PM11H12G	Penicilli
29	62	4.9	136551	2	AC048354	Homo sapi
30	60.8	4.8	160327	4	AC149763	Bos tauru
31	60.8	4.8	240521	2	AC149759	Bos tauru
32	60.4	4.7	155337	2	AC116408	Mus muscu
33	58.6	4.6	155337	2	AC116408	Mus muscu
34	58.6	4.6	202851	10	AC115880	Mus muscu
35	58.4	4.6	248550	1	SCO939120	Streptomy
36	58.2	4.6	94058	2	AC141024	Rattus no
37	58	4.5	92310	2	AC108380	Pan trogl
38	57.8	4.5	2000	6	AX655393	Sequence
39	57.2	4.5	143406	10	AC092857	Rattus no
40	57.2	4.5	230329	2	AC142070	Rattus no
41	56.6	4.4	245210	2	AC137771	Homo sapi
42	56.2	4.4	326244	2	AC118627	Mus muscu
43	56	4.4	142402	2	CR407549	Danio rer
44	56	4.4	247644	5	BX470223	Zebrafish
45	56	4.4	250011	5	BX548044	Zebrafish

ALIGNMENTS

RESULT 1	AY191415	1260 bp	mrna	linear	ROD 12-MAR-2003
LOCUS	AY191415				
DEFINITION	Mus musculus zygote arrest 1 (Zarl) mRNA, complete cds.				
ACCESSION	AY191415				
VERSION	AY191415.1	GI:27808689			
KEYWORDS	Mus musculus (house mouse)				
SOURCE	Mus musculus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 1260)				
AUTHORS	Wu, X., Viveiros, M.M., Eppig, J.J., Bai, Y., Fitzpatrick, S.L. and Matzuk, M.M.				
TITLE	Zygote arrest 1 (Zarl) is a novel maternal-effect gene critical for the oocyte-to-embryo transition				
JOURNAL	Nat. Genet. 33 (2), 187-191 (2003)				
MEDLINE	22447938				
PUBMED	12539046				
REFERENCE	2 (bases 1 to 1260)				
AUTHORS	Wu, X., Wang, P. and Matzuk, M.M.				
TITLE	Direct Submision				
JOURNAL	Submitted (04-DEC-2002) Pathology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA				
FEATURES	Location/Qualifiers				
source	1..1260				
	/organism="Mus musculus"				
	/mol_type="mRNA"				
	/db_xref="taxon:10090"				
	/chromosome="5"				
	/map="between DSBuc48 and Ttk"				
gene	1..1260				
	/gene="Zarl"				
5'UTR	1..27				
	/gene="Zarl"				
CDS	28..1113				
	/gene="Zarl"				
	/note="Zarl; maternal factor; contains atypical C8 PHD motif at C-terminus"				
	/codon_start=1				
	/product="zygote arrest 1"				

/protein_id="AA024706.1"
 /db_xref="GI:27808690"
 /translation="MFPASTPHPCPHVPOATKAGDWRGARGCRPAAPSPFLPGYRQ
 LMAAYVDHSHORQAQIMALLSRMGPRSVSSRDAAVQVNPDRDASVQCSLSGRITLQPGC
 RASPDARSQCPQHRGAGAPRSWQTVAPFSSVTFGLSSSLEVAGGQTPTKGBG
 SPASGSTRPSPRAAKAVQPRSEEDVQAAAGWQPPPPDRNSVAAWOSFP
 GSEPPCAAEQAQPDQASQSTEQDKERLRFQLEQYGYHCKDKIRW
 ESAYVWCQGTSKVYFQFCRVCEKSNPYRVEDITCQSKRTRCACPVLRLHVDPR
 PHRDLCGRCKDKRLSDSTFSFKYII"
 1114..1260
 3'UTR
 /gene="Zarl"
 ORIGIN
 Query Match 97.5%; Score 1245.4; DB 10; Length 1260;
 Best Local Similarity 99.8%; Pred. No. 2,1e-248;
 Matches 1257; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY	3	GGCGGGCGAGGCGCGGACGCA	CCCATGTTCCCGCGGAGCGCTTCCACCCCTGCCCGCA	62
DB	3	GGCGGGCGAGGCGCGGACGCA	CCCATGTTCCCGCGGAGCGCTTCCACCCCTGCCCGCA	62
QY	63	TCCTTTATCGCAGGCCACCA	AGCGGGATGGCTGGAGGTTCCGAGCAGGGGCTGCCG	122
DB	63	TCCTTTATCGCAGGCCACCA	AGCGGGATGGCTGGAGGTTCCGAGCAGGGGCTGCCG	122
QY	123	ACCGCGCCCCCTCTCTCT	CCCCGGCTACAGACAGCTCATGGCCGCGGAGTACGTGCA	182
DB	123	ACCGCGCCCCCTCTCTCT	CCCCGGCTACAGACAGCTCATGGCCGCGGAGTACGTGCA	182
QY	183	CAGCCACCAAGCGGGCACA	CTCATGGCCCTGTGTGGGATGGTCCCGTCCGTCAG	242
DB	183	CAGCCACCAAGCGGGCACA	CTCATGGCCCTGTGTGGGATGGTCCCGTCCGTCAG	242
QY	243	CAGCGGTGACGTGGCGT	GCAGGTGAACCCGCGCGGACGCTCGGTGACGTGTTCACT	302
DB	243	CAGCGGTGACGTGGCGT	GCAGGTGAACCCGCGCGGACGCTCGGTGACGTGTTCACT	302
QY	303	CGGGCGCGCAGCTGCAG	CGCTGAGGGTCCGAGCCCGGACCGCCCGATCGGGTTC	362
DB	303	CGGGCGCGCAGCTGCAG	CGCTGAGGGTCCGAGCCCGGACCGCCCGATCGGGTTC	362
QY	363	CTGTCAACCCGNGCCAG	CGCGCGCGGAGATCCCGCGATCTCTGACAGACCGTAGC	422
DB	363	CTGTCAACCCGNGCCAG	CGCGCGCGGAGATCCCGCGATCTCTGACAGACCGTAGC	422
QY	423	CCCGTTCTCGTCTGAC	CTTCTGTGGCTCTCTCTCACTGGAGTTGCGGGAGGCAG	482
DB	423	CCCGTTCTCGTCTGAC	CTTCTGTGGCTCTCTCTCACTGGAGTTGCGGGAGGCAG	482
QY	483	GCAGACACCAAGAGG	AGAGGGAGCCCGGATCTCTGGGACCCGGGACCCGAGCC	542
DB	483	GCAGACACCAAGAGG	AGAGGGAGCCCGGATCTCTGGGACCCGGGACCCGAGCC	542
QY	543	GAGAGAGTGGCGCG	AGAGAAACGCTCCCGACCGCGAGCGAGGGCGATTTCA	602
DB	543	GAGAGAGTGGCGCG	AGAGAAACGCTCCCGACCGCGAGCGAGGGCGATTTCA	602
QY	603	GGCTGACGGCAGCG	CGGGTGGAGCAGCACCACCGAGGACCGGAACAGTGTGC	662
DB	603	GGCTGACGGCAGCG	CGGGTGGAGCAGCACCACCGAGGACCGGAACAGTGTGC	662
QY	663	GGCGATGACGTGAG	CTGGAGCGGAGGACCATGTCCTGCCGAGAGATGCTCAGGA	722
DB	663	GGCGATGACGTGAG	CTGGAGCGGAGGACCATGTCCTGCCGAGAGATGCTCAGGA	722
QY	723	CCCCGGTGATTCGAT	CCCCCTCCAGACCGAGGCTCCCGCAAGCAGCGAGCAGACAA	782
DB	723	CCCCGGTGATTCGAT	CCCCCTCCAGACCGAGGCTCCCGCAAGCAGCGAGCAGACAA	782
QY	783	GGAGCGCTGGTTTCA	GTTTCTTAGACGAGAGTAGGCTACTATCACTGCAAGGACTG	842
DB	783	GGAGCGCTGGTTTCA	GTTTCTTAGACGAGAGTAGGCTACTATCACTGCAAGGACTG	842

QY	843	CAAAATCCCGTGGGAGAGCGCC	TATGTGTGTGTGTGTCAGGGCACCAGTAAGGTGTACT	902
DB	843	CAAAATCCCGTGGGAGAGCGCC	TATGTGTGTGTGTGTCAGGGCACCAGTAAGGTGTACT	901
QY	903	TCAAAACAGTTCTCCGAG	TGTGTGAGAAATCTTCAACCCCTTACAGAGTGGAGGACATCA	962
DB	902	TCAAAACAGTTCTCCGAG	TGTGTGAGAAATCTTCAACCCCTTACAGAGTGGAGGACATCA	961
QY	963	CTGTCAAAGTTGTAAAGAA	CTAGATGTGCTGCCAGTCAGATTTGCCACGTGGACC	1022
DB	962	CTGTCAAAGTTGTAAAGAA	CTAGATGTGCTGCCAGTCAGATTTGCCACGTGGACC	1021
QY	1023	CTAAACGCCCCCATCGCA	AGACTTGTGTGGAGATGCAAGCAAAACGCTGTCTCTGCG	1082
DB	1022	CTAAACGCCCCCATCGCA	AGACTTGTGTGGAGATGCAAGCAAAACGCTGTCTCTGCG	1081
QY	1083	ACAGCACCTTCAGCTTCA	AAATACATCATTTAGTGAGAGTCAAAACGTTTCTGCTAGATG	1142
DB	1082	ACAGCACCTTCAGCTTCA	AAATACATCATTTAGTGAGAGTCAAAACGTTTCTGCTAGATG	1141
QY	1143	GGGCTAATGGAATGGACA	GTGAGCTTCTCCCTTCCCTTCCCAATTC	1202
DB	1142	GGGCTAATGGAATGGACA	GTGAGCTTCTCCCTTCCCTTCCCAATTC	1201
QY	1203	TTCATGACAGACAGTGT	TACTTGGATATAAGCTGTGAATAAAAGTATTGCAAAACA	1261
DB	1202	TTCATGACAGACAGTGT	TACTTGGATATAAGCTGTGAATAAAAGTATTGCAAAACA	1260

RESULT 2
 AV283175 1280 bp mRNA linear ROD 21-AUG-2003
 LOCUS Rattus norvegicus zygote arrest 1 (Zarl) mRNA, complete cds.
 DEFINITION
 ACCESSION AV283175
 VERSION AV283175.1 GI:30908932
 KEYWORDS
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 1280)
 AUTHORS Wu, X., Wang, P., Brown, C.A., Zilinski, C.A. and Matzuk, M.M.
 TITLE Zygote arrest 1 (zar1) is an evolutionarily conserved gene expressed in vertebrate ovaries
 JOURNAL Biol. Reprod. 69 (3), 861-867 (2003)
 MEDLINE 22811438
 PUBMED 12773403
 RREFERENCE 2 (bases 1 to 1280)
 AUTHORS Wu, X. and Matzuk, M.M.
 TITLE Direct Submission
 JOURNAL Submitted (24-APR-2003) Pathology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

FEATURES
 source
 1..1280
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /chromosome="14"
 1..1280
 /gene="Zarl"
 28..1113
 /gene="Zarl"
 /note="maternal factor"
 /codon_start=1
 /product="zygote arrest 1"
 /protein_id="AAP37037.1"
 /db_xref="GI:30908933"
 /translation="MFPASTPHPCPHVPOATKAGDWRGARGCRPAAPSPFLPGYR
 QLMAYYDVSQRAQIMALLSRMGPRSVSSRDAAVQVNPDRDASVQCSLSGRITLQPGR
 RASPDARSQCPQHRGAGAPRSWQTVAPFSSVTFGLSSSLEVAGGQTPTKGBG
 GSEPPCAAEQAQPDQASQSTEQDKERLRFQLEQYGYHCKDKIRW
 ESAYVWCQGTSKVYFQFCRVCEKSNPYRVEDITCQSKRTRCACPVLRLHVDPR
 PHRDLCGRCKDKRLSDSTFSFKYII"

ESAYWVCVQGTSKVYKQFCRVCKSYNRYVEDITQSCKRTRCAPVRLRHVDPKR
PHRDLOGRCKDKLSDFSFKYII"

ORIGIN

Query Match 75.3%; Score 962.2; DB 10; Length 1280;
Best Local Similarity 87.7%; Pred. No. 1.5e-189;
Matches 1110; Conservative 0; Mismatches 143; Indels 12; Gaps 5;

Qy	3	GGCGGGGAGCGCGGGAGCGACCATGTTCCCGCGAGCAGCTTCCACCCCTGCGCGCA	62
Db	3	GGCGGGGAGCGCGGGAGCGACCATGTTCCCGCGAGCAGCGCCACCATGCGCGCA	62
Qy	63	TCCTTATCCGC---AGGCCACAAAGCCGGGATGCTGGAGGTCGGAGCAGGGGCTG	119
Db	63	TCCTTATCCGC---AGGCCACAAAGCCGGGATGCTGGAGGTCGGAGCAGGGGCTG	122
Qy	120	CGGACCGCGCGCCCTCTTCTCCCGGCTACAGACAGCTCATGGCCGGGAGTAGCT	179
Db	123	CAGCCCGGAGCCCTCTTCTCCCGGCTACAGACAGCTCATGGCCGGGAGTAGCT	182
Qy	180	CGACAGCCAGCGGGGACAGCTCATGGCCCTGCTGCGCGGATGGGTCCCGGCTGGT	239
Db	183	TGACAGCTATCAGCGGGGAGCTCATGGCCCTGCTGCGCGGATGGGTCCCGGCTGGT	242
Qy	240	CAGCAGCGTGAAGTGGGTGAGGTAACCCCGCGCGAGATCCCGGCTGCTGCTGCTTC	299
Db	243	CAGCAGCGGACGCTGCGGTGAGGTAACCCCGCGCGAGATCCCGGCTGCTGCTGCTTC	302
Qy	300	ACTGGGGCGCGAGCTGAGCTGAGGCTGCGGAGTGGGAGCGCCGCGGAGTGGG	359
Db	303	GCTGGGGCGCGACACTGAGCTGAGGCTGCGGAGTGGGAGCGCCGCGGAGTGGG	362
Qy	360	TTCTGTCAACCCCGTGGCGAGCGCGCGCGGAGATCCCGGCTGCTGCGAGACCGT	419
Db	363	TTCTGTCAACCCCGTGGCGAGCGCGCGGAGATCCCGGCTGCTGCGAGACCGT	422
Qy	420	AGCCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	479
Db	423	CGCCCTGTACTCGCCCGTACCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	482
Qy	480	CAGCAGACACCCAGAGGAGAGGAGCGCGCGGATCCCTGCGGAGCGCGGAGCGGA	539
Db	483	CAGCAGACACCCAGAGGAGAGGAGCGCGCGGATCCCTGCGGAGCGCGGAGCGGA	542
Qy	540	GCCGAGAGAGTGGCGCGGAGGAGGAGCGCGCGGATCCCTGCGGAGCGCGGAGCG	599
Db	543	GCCGAGAGAGTGGCGCGGAGGAGGAGCGCGCGGATCCCTGCGGAGCGCGGAGCG	602
Qy	600	TCAGGCTGAGGGGAGGCGCGGAGGAGGAGCGCGCGGATCCCTGCGGAGCGCGGAG	659
Db	603	CCAGGCTGAGGGGAGGCGCGGAGGAGGAGCGCGCGGATCCCTGCGGAGCGCGGAG	659
Qy	660	GGCGCGGATGAGTGTGAGCTGGAGGAGGAGGAGCGCGCGGATCCCTGCGGAGCG	719
Db	660	GGCGCGGATGAGTGTGAGCTGGAGGAGGAGGAGCGCGCGGATCCCTGCGGAGCG	719
Qy	720	GGACCCCGTGTATTCGAGTGGCTTCGAGAGCGCGGATCCCTGCGGAGCGCGGAG	779
Db	720	GGACCCCGTGTATTCGAGTGGCTTCGAGAGCGCGGATCCCTGCGGAGCGCGGAG	779
Qy	780	CAAGGAGCGCGTGGCTTTCCAGTTCTTTAGAGCAGAGTACGGCTACTATCTCAAGGA	839
Db	780	CAAGGAGCGCGTGGCTTTCCAGTTCTTTAGAGCAGAGTACGGCTACTATCTCAAGGA	839
Qy	840	CTGCAAAATCCGTTGGAGAGCGCTTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGT	899
Db	840	CTGCAAAATCCGTTGGAGAGCGCTTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGT	898
Qy	900	ACTTCAAAACAGTCTTCCGCGAGTGTGTGAGAAATCTTCAACCCCTTACAGAGTGGAGCA	959
Db	899	ACTTCAAAACAGTCTTCCGCGAGTGTGTGAGAAATCTTCAACCCCTTACAGAGTGGAGTA	958
Qy	960	TCACCTGTCAAAGTTGTAAAGAACTAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1019

Db	959	TCACCTGCCAAAGTTGTAAAGGACTAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1018
Qy	1020	ACCTAAACGCCCGCATCGGCAAGACTTGTGTGGAGATGCAAGGCAAAACGCTGTCTCT	1079
Db	1019	ACCTAAACGCCCGCATCGTCAAGACTTGTGTGGAGATGCAAGGCAAAACGCTGTCTCT	1078
Qy	1080	GCAGCAGCAGCTTCAGCTTCAATACATCATTTAGTGAAGTTCGAAAACGTTTCTGCTA-	1138
Db	1079	GTGACAGCAGCTTCAGCTTCAATATATCATTTAGAGAGGTTAAATATGTTTCTGCTAA	1138
Qy	1139	---GATGGGGCTAATGGAATGGAAGTGAAGTCTTCTCCCTCTTTCACCTCTTCCCTTTC	1195
Db	1139	ATGGATCGGA CAAGTGAGCCTTCTCCAGCGCCCTCTCCCTCTCTCCCTCTCTCCCTCTC	1198
Qy	1196	CAAAATCTTCATGACAGACAGTGTACTTGGATATAAGCTGTGAATAAAGGATATTGC	1255
Db	1199	AAAATACTTTCATGAAGGCGAGTG-TACTTTAATAAAGCTGCGAATAAAGGGGTTGC	1257
Qy	1256	AAACA 1260	
Db	1258	AAACA 1262	
RESULT 3			
LOCUS AV193889 3987 bp DNA linear ROD 12-MAR-2003			
DEFINITION Mus musculus zygote arrest 1 (Zarl) gene, complete cds.			
ACCESSION AV193889			
VERSION AV193889.2 GI:27808699			
KEYWORDS Mus musculus (house mouse)			
ORGANISM Mus musculus			
REFERENCE 1 (bases 1 to 3987)			
AUTHORS Wu, X., Viveiros, M.M., Eppig, J.J., Bai, Y., Fitzpatrick, S.L. and Matzuk, M.M.			
TITLE Zygote arrest 1 (Zarl) is a novel maternal-effect gene critical for the oocyte-to-embryo transition			
JOURNAL Nat. Genet. 33 (2), 187-191 (2003)			
MEDLINE 22447938			
PUBMED 12539046			
REFERENCE 2 (bases 1 to 3987)			
AUTHORS Wu, X. and Matzuk, M.M.			
TITLE Direct Submission			
JOURNAL Submitted (09-DEC-2002) Pathology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA			
REFERENCE 3 (bases 1 to 3987)			
AUTHORS Wu, X. and Matzuk, M.M.			
TITLE Direct Submission			
JOURNAL Submitted (21-JAN-2003) Pathology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA			
REMARK Sequence update by submitter			
COMMENT On Jan 21, 2003 this sequence version replaced gi:27808693.			
FEATURES			
source			
1. .3987			
/organism="Mus musculus"			
/mol_type="genomic DNA"			
/strain="129/SvEv"			
/db_xref="taxon:10090"			
1. .3987			
/gene="Zarl"			
join(1..801,2574..2666,2745..2819,3697..3987)			
/gene="Zarl"			
/product="zygote arrest 1"			
1. .801			
/gene="Zarl"			
/number=1			
join(28..801,2574..2666,2745..2819,3697..3987)			
/gene="Zarl"			
/note="oocyte-specific"			
/codon_start=1			
gene			
mRNA			
exon			
CDS			

		/product="zygote arrest 1"	
		/protein_id="AA024708.1"	
		/db_xref="GI:27808694"	
		/translations="MPPASTFPHCPHYQATKAGDWRFGARCFARPSPFLEGVYQ LMAEYVDHSORAOILMALLSRMGPPRSVSROAAQVNPNRDASVQCSLGRRTLPQAGC RASPDDARSCOPRGHAGKAGRSPRSMOTVAPSSVTCGLSSLSLEAVAGRGQTTKGEG SPASGTRREPREPREVAARAGAVQPPRSEGDVQAAGWEOQPPPEDRNSVAMQSEP GSEBPCPAEMADQDQSDAPRDQASPOSTEQDKERLARFQLEQKYGVYHCKDKLIRW ESAYVVCVQSTSYVYKQFCRVCESKSYNPIRVEDITTCOSCKRKTCAFCVRLRHWDPKR FHRQDLGRCRCKRLLSCDSFTFSFKYII"	
exon		/gene="Zar1"	2574. .2666
		/number=2	
exon		/gene="Zar1"	2745. .2819
		/number=3	
exon		/gene="Zar1"	3697. .3987
		/number=4	
ORIGIN			
Query Match		62.6%;	Score 799.4; DB 10; Length 3987;
Best Local Similarity		99.9%;	Pred. No. 1e-155; 1; Indels 0; Gaps 0;
Matches 800; Conservative 0;		Mismatches	
QY	3.	GGCGGGCAGAGCGGGGAGCAGCACCATATGTTCCGCGGAGACAGCTTCCACCCCTGCCGCA	62
DB	3	GGCGGGCAGAGCGGGGAGCAGCACCATATGTTCCGCGGAGACAGCTTCCACCCCTGCCGCA	62
QY	63	TCCTTTATCGCAGGCCACCAAGCCGGGATGCTCGAGGTTTCGGAGCCAGAGGGCTGC	122
DB	63	TCCTTTATCGCAGGCCACCAAGCCGGGATGCTCGAGGTTTCGGAGCCAGAGGGCTGC	122
QY	123	ACCCGCGCCCCCTCTCTTCTCCCGCGCTACAGACAGCTCATGCGCGCGGAGTACGTGCA	182
DB	123	ACCCGCGCCCCCTCTCTTCTCCCGCGCTACAGACAGCTCATGCGCGCGGAGTACGTGCA	182
QY	183	CAGCCACAGCGGGCAGACGCTCATGCGCCCTGCTGTGCGGATGGGTCCCGGTTCGGTCAG	242
DB	183	CAGCCACAGCGGGCAGACGCTCATGCGCCCTGCTGTGCGGATGGGTCCCGGTTCGGTCAG	242
QY	243	CAGCCGTGACGTGCGGTGCAGGTGAACCCGCGCGCGAGTCACTTCGGTGAGTGTCACT	302
DB	243	CAGCCGTGACGTGCGGTGCAGGTGAACCCGCGCGCGAGTCACTTCGGTGAGTGTCACT	302
QY	303	CGGGCCCGCAGCTGTGAGCTTCAGAGGTGCGGAGCCAGCCCGCGATCGGGTTC	362
DB	303	CGGGCCCGCAGCTGTGAGCTTCAGAGGTGCGGAGCCAGCCCGCGATCGGGTTC	362
QY	363	CTGTCAACCCGTGGGCCACCGCGCGCGCGGAGATCCCGCGATCTCTGGCAGACCGTAGC	422
DB	363	CTGTCAACCCGTGGGCCACCGCGCGCGCGGAGATCCCGCGATCTCTGGCAGACCGTAGC	422
QY	423	CCCGTTCTCGTCGTGACCTTCTGTGGCTCTCTCTCTCACTCGAGGTTTCGGGAGCGAG	482
DB	423	CCCGTTCTCGTCGTGACCTTCTGTGGCTCTCTCTCTCACTCGAGGTTTCGGGAGCGAG	482
QY	483	GCAGACACCCAGAGGGAGAGGGGAGCCCGGCATCTCTCGGGACCCGGGAACCGGAGCC	542
DB	483	GCAGACACCCAGAGGGAGAGGGGAGCCCGGCATCTCTCGGGACCCGGGAACCGGAGCC	542
QY	543	GAGAGAGGTGGCCGAGGAAAGCGGTTCCTCCAGCCGGGAGCGAGAGGCGCATGTTC	602
DB	543	GAGAGAGGTGGCCGAGGAAAGCGGTTCCTCCAGCCGGGAGCGAGAGGCGCATGTTC	602
QY	603	GGCTGCAGGGCAGGCCGGGTGGGAGCAGCAGCCACCCGAGGAGCCGGAACCGTAGTGGC	662
DB	603	GGCTGCAGGGCAGGCCGGGTGGGAGCAGCAGCCACCCGAGGAGCCGGAACCGTAGTGGC	662
QY	663	GGCGATGCACTGTGACCTGGGAGCGAGGCCATGTCTCTCCGACAGAGATGGCTCAGGA	722
DB	663	GGCGATGCACTGTGACCTGGGAGCGAGGCCATGTCTCTCCGACAGAGTGGCTCAGGA	722

Qy	723	CCCGGTGATTGGATGCGCTTCGACGACGAGGCTCCCGCGAAAGCAGCGAGGACGACAA	782
Db	723	CCCGGTGATTGGATGCGCTTCGACGACGAGGCTCCCGCGAAAGCAGCGAGGACGACAA	782
Qy	783	GGAGCGCGCTGCGGTTTCCAGTT	803
Db	783	GGAGCGCGCTGCGGTTTCCAGTT	803
RESULT 4			
AC127733			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			

TITLE Direct Submission
JOURNAL Submitted (08-AUG-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Jul 4, 2004 this sequence version replaced gi:46240945.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RN/RepeatMasker.html>
 ----- Genome Center
 Center: Whitehead Institute/MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@broad.mit.edu
 ----- Project Information
 Center project name: L17506
 Center clone name: 506 B 15

```

ORIGIN
/clone_11b="RPci-24 Male Mouse BAC"

Query Match      62.6%; Score 799.4; DB 2; Length 144954;
Best Local Similarity 99.9%; Pred.No. 9.9e-156;
Matches 800; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GGCGGGCGAGCGCGGGGACGACCCATGTTCCCGGGCGAGCACGTTTCAACCCCTGCCGCGCA 62

```

Db	110537	GGCGGCGAGCGCGGACGACCAATGTTCCCGCGAGCAGTTCCACCCCTGCCCGCA	110596
Qy	63	TCCTTATCCGAGGACCAAAAGCCGGGATGGCTGGAGTTCCGAGCAGGGGCTGCCG	122
Db	110597	TCCTTATCCGAGGCGACCAAAAGCCGGGATGGCTGGAGTTCCGAGCAGGGGCTGCCG	110656
Qy	123	ACCCGCGCCCTCTCTTCCTCCCGGTACAGACAGCTCATGCGCGGAGTACGTGCA	182
Db	110657	ACCCGCGCCCTCTCTTCCTCCCGGTACAGACAGCTCATGCGCGGAGTACGTGCA	110716
Qy	183	CAGCCACGAGGGGACAGCTCATGGCCCTGCTGTGCGGATGGTCCCGGTCCGTCAG	242
Db	110717	CAGCCACGAGGGGACAGCTCATGGCCCTGCTGTGCGGATGGTCCCGGTCCGTCAG	110776
Qy	243	CAGCCGTGACGCTCGGTTGCAGGTGAACCCGCGCGGACGCCCTCGGTGCAGTTCCT	302
Db	110777	CAGCCGTGACGCTCGGTTGCAGGTGAACCCGCGCGGACGCCCTCGGTGCAGTTCCT	110836
Qy	303	CGGCGCGGACGCTGCAGGCTGCGAGCGGAGCCAGCCGCCCGATCGGGTTC	362
Db	110837	CGGCGCGGACGCTGCAGGCTGCGAGCGGAGCCAGCCGCCCGATCGGGTTC	110896
Qy	363	CTGTCAACCCCGTGGCCACGCGCGCGCGGAGATCCCGCGATCTTGGCAGACCGTAGC	422
Db	110897	CTGTCAACCCCGTGGCCACGCGCGCGCGGAGATCCCGCGATCTTGGCAGACCGTAGC	110956
Qy	423	CCCGTTCTCGTCCGTGACCTTCTGTGGCTCTCTCTCTCACTGAGGTTGCGGAGGCGAG	482
Db	110957	CCCGTTCTCGTCCGTGACCTTCTGTGGCTCTCTCTCTCACTGAGGTTGCGGAGGCGAG	111016
Qy	483	GCAGACACCCAGAGGAGGAGGCGGAGCCCGGCATCTCTCGGGACCCGGGAACCGGAGCC	542
Db	111017	GCAGACACCCAGAGGAGGAGGCGGAGCCCGGCATCTCTCGGGACCCGGGAACCGGAGCC	111076
Qy	543	GAGAGAGGTGACCGCGAGGAACCGGTCCCCCAGCCCGCGAAGCAGGAGGCGCATGTTCA	602
Db	111077	GAGAGAGGTGACCGCGAGGAACCGGTCCCCCAGCCCGCGAAGCAGGAGGCGCATGTTCA	111136
Qy	603	GGCTGACGGGACGCGGGTGGAGACAGACCAACCCGAGAGCCGGAACAGTGTGGC	662
Db	111137	GGCTGACGGGACGCGGGTGGAGACAGACCAACCCGAGAGCCGGAACAGTGTGGC	111196
Qy	663	GGCGATCAGTCTGAGCTGSGAGCGAGGAGCCATGTCTCCGACAGATGGCTCAGGA	722
Db	111197	GGCGATCAGTCTGAGCTGSGAGCGAGGAGCCATGTCTCCGACAGATGGCTCAGGA	111256
Qy	723	CCCGGTGATTCCGATGCCCTCCGAGACCCAGGCTCCCGCAAGCAGCGAGCAGCA	782
Db	111257	CCCGGTGATTCCGATGCCCTCCGAGACCCAGGCTCCCGCAAGCAGCGAGCAGCA	111316
Qy	783	GGAGCGCTCGGTTTCAGTT	803
Db	111317	GGAGCGCTCGGTTTCAGTT	111337

RESULT 5	
AL671880	200535 bp DNA linear ROD 27-NOV-2002
LOCUS	Mouse DNA sequence from clone Rp23-384C22 on chromosome X, complete
DEFINITION	sequence.
ACCESSION	AL671880
VERSION	AL671880
KEYWORDS	HTG.
SOURCE	AL671890.15 GI:25896274
ORGANISM	Mus musculus (house mouse)
	Mus musculus
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 200535)
AUTHORS	Van Hellmond,Z.
TITLE	Direct Submission
JOURNAL	Submitted (27-NOV-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GZCZ
Center clone name: CH230-159N5
----- Summary Statistics

```

Assembly program: Phrap; version 0.990329
 Consensus quality: 216235 bases at least Q40
 Consensus quality: 218057 bases at least Q30
 Consensus quality: 219087 bases at least Q20
 Consensus quality: 219408 bases at least Q20
 Estimated insert size: 221324; sum-of-configs
 quality coverage: 7x in Q20 bases sum-of-con-

```

Quality coverage: 7X in Q20 bases; sum-of-contigs coverage:
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html).
*
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
*
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 235390: contig of 235390 bp in length.

```

```

FEATURES
  source
    1. .235390
      Location/Qualifiers
        /organism="Rattus norvegicus"
        /mol_type="genomic DNA"
        /db_xref="taxon:10116"
        /align="CH230-159N5"

```

```
misc_feature
1. .1856
/note="wgs_end_extension
clone_end:77"
misc_feature
complement(4755..38027)
/note="clone_boundary
```

```
misc_feature
clone_end:17
site:
end_sequence:BH321013"
complement(227221..228048)
/note="clone_boundary
clone_end:Sp6
```

```
misc feature
site:
end sequence:BH321015"
232121. .235390
/note="wgs_end_extension
clone end:Sp6"
```

ORIGIN

Query Match	49.2%;	Score 628; DB 2;	Length 235390;
Best Local Similarity	88.1%;	Pred. No. 4.1e-120;	
Matches 708. Conservative		0: Mismatches 90;	Indels 6;

QY	3	GGCGGGCGAGCGCGGCGCGACCCCATCTTCGCGGCGAGCAGTTTCCACCCCTGCCGCGCA	62
MATCHES			
706;	Conservative		

DB		27469	GGCGGCGAGGCCGGAGACGCACCATCTGTCCTCCGCGAGCAGCAAGTGTCTTTTG		119
QY	63	TCCTTATCCGC-----AGGCCACCAAAGCCGGGGATGCTTGGAGGTTCCGAGCCACGGSGTG			119

Db	27409 TCCTTACCGCCCAACGGCAGCCAAAGCCGCGGATGGCTCGAGATTTCGAGCACCGAGGGGCTG 21355
Oy	120 CCGAACCGGCGCCCCCTCTTTCTCCCCGGCTACAGACAGCTCATGGCCCGCAGTAGCT 179

Db	27349	CAGGCCCCGAGGCCCCCTCTTCTCTCCCGCTACAGACAGCTCATGGCCCGGAGTACTT	27299
QY	180	CGACAGCCACGAGGGGNCACGCTCATGGCCCTGCTCTCGCGATGGTCCCGTCCGT	239

[illegible]

Unpublished
2 (bases 1 to 235390)
Worley, K.C.

Direct Submission
Submitted (06-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 235390)

Rat Genome Sequencing Consortium.

Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On Nov 20, 2002 this sequence version replaced gi:23195602. The sequence in this assembly is a combination of BAC reads from the whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence

may extend beyond the ends of the stone and


```

Consensus quality: 232679 bases at least Q20
Estimated insert size: 242796; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

-----
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  consists of 1 contigs. Gaps between the contigs
  are represented as runs of N. The order of the pieces
  is believed to be correct as given, however the sizes
  of the gaps between them are based on estimates that have
  * provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved..
* 1 237695: contig of 237655 bp in length.

FEATURES             Location/Qualifiers
     source            1..237695
                        organism="Rattus norvegicus"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:10116"
                        /clone="CH230-74L11"
     misc_feature      1..1905
                        note="wgs end_extension
                        clone_end:T7"
     misc_feature      complement(5394..6172)
                        note="clone_boundary
                        clone_end:T7"
     misc_feature      site:EcoRI
                        end_sequence:BH339391"
                        complement(234275..235086)
                        note="clone_boundary
                        clone_end:Sp6
                        site:EcoRI
                        end_sequence:BH339393"

ORIGIN
Query Match          49.2%; Score 628; DB 2; Length 237695;
Best Local Similarity 88.1%; Pred. No. 4.le-140; Indels   6; Gaps    2;
Matches 708; Conservative 0; Mismatches 90;

QY  3  GCCGGGCGAGCGCGGGACGCACCATGTTCCTCCGCCGAGCAGTTCCACCCCTGCCGCCA  62
Db   123013  GCCGGGCGAGCGCGGGACGCACCATGTTCCTCCGCCGAGCAGCCCCACCATGCCCA  122954

QY  63  TCCTTATCCGC---AGGCCAACAAAGCGGGGATGGCTGGAGGTTCCGAGCCAGGGCTG  119
Db   122953  TCCTTATCCGC---AGGCCAACAAAGCGGGGATGGCTGGAGGTTTGAGCCAGGGCTG  122894

QY  120  CCGACCCGCGCCCCCTCTTCTCTCCCGGCTACAGACAGCTCATGGCCGCGGAGTAGCT  179
Db   122893  CAGGCCGAGCCCCCTCTTCTCTCCCGGCTACAGACAGCTCATGGCCGCGGAGTAGCT  122834

QY  180  CGACAGCCACAGCGGGGACAGCTCATGGCCCTGCTGTCGGGATGGTCCCCTGCGGT  239
Db   122833  TCACAGCTATTACGAGAGCGCAGCTCATGGCCCTGCTGTCGGGATGGTCCCCTGCGGT  122774

QY  240  CAGCAGCCGTACAGCTGCGGTGACAGTGAACCCGCGCGCGAGATCCCGCTCGGTGCTTC  299
Db   122773  CAGCAGCCGCGAGCTGCGGTGACAGTGAACCCGCGCGCGAGATCCCGCTCGGTGCTTC  122714

QY  300  ACTGGGGCGCGCACGCTGACGCTGAGGGTCCGAGCCAGCCCGCCGAGCCCGATCCGG  359
Db   122713  GCTCGGGGCGCGCACACTGACGCTGAGCGGGCCGAGCCAGCCCGCCGAGCCCGATCCGG  122654

QY  360  TTCCTGTCAACCCCGTGGCCACGCGCGGCGCCGGGAGATCCCCCGCATCTTGCAGACCG  419
Db   122653  TTCCTGTCAACCCCGGAGCCGCGCGGCGCCGGGAGATCCCCCGCATCTTGCAGACCG  122594

QY  420  AGCCCGGTTCCTCGTGGTGAACCTTGTGGCTCTCTCTCTCACTGGAGGTTCCGGGAG  479
Db   122593  CGCCCTGTACTCGCGGTGAACCTTGGTGGCTCTCTCTCTCACTGGAGGTTCCGGGAG  122534

```


All repeats were identified using RepeatMasker:
Smt, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www.seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

AC108848/C	AC108848	220894 bp	DNA	linear	HTG 13-MAR-2004
LOCUS	Mus musculus	chromosome 5	clone	RP23-300K5	map 5, WORKING DRAFT
DEFINITION	SEQUENCE, 40	unordered pieces.			

AC108848	
AC108849	
AC108848.3	GI.45430126
HTG; HTGS PHASE1; HTGS DRAFT;	HTGS_FULLTOP.
Mus musculus (house mouse)	
Mus musculus	
ORGANISM	
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
1 (bases 1 to 220894)	
REFERENCE	

AUTHORS Birren, B., Nusbaum, C. and Lander, E.
TITLE Mus musculus chromosome 5, clone RP23-300K5
JOURNAL Unpublished

2 (bases 1 to 220894), Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Barren, B., Linton, L., Bastien, V., Boguslavsky, L., Boukagalter, B.,
Anderson, S., Barna, N., Campione, A., Chang, J., Chazaro, B.,
Brown, A., Camarata, J., Collins, S., Collamore, A., Cook, A.,
Choepei, Y., Colangelo, M., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
Cooke, P., Dearellano, K., Gage, D., Galagan, J., Gardyna, S.,
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Gins, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,
Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,
Landers, T., Lehotsky, J., Levine, R., Liu, G., MacLean, C.,
Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
McSwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T.,
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, P., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
Robetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S.,
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Traversman, M., Travis, N., Triggilo, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
Direct Submission
JOURNAL
Submitted (31-JAN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE
3 (bases 1 to 220894)

REFERENCE
AUTHORS

Birzen, B., Nubaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H. M., Barna, N., Baatien, V., Bloom, T., Boguslavsky, L., Boukgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corriu, B., DeArellano, K., Dias, J. S., Dodge, S., Dooley, K., Corris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Ton, K., Liu, X., Lui, A., Mabbitt, R., MacLean, C., McDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Melgrim, J., Meneses, L., Mihova, T., Mithwa, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Teafaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

TITLE

JOURNAL	COMMENT
Submitted (13-MAR-2004)	Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 13, 2004	this sequence version replaced gi:20336129.

THE UNIVERSITY OF CHICAGO

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2005, 14:32:42 ; Search time 679 Seconds
(without alignments)
9872.622 Million cell updates/sec

Title: US-09-830-810A-1

Perfect score: 1277

Sequence: 1 aagcgggcgagcgcgga.....acaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_23Sep04:*

1: geneseqn1980s:*

2: geneseqn1990s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002as:*

7: geneseqn2002bs:*

8: geneseqn2003as:*

9: geneseqn2003bs:*

10: geneseqn2003cs:*

11: geneseqn2003ds:*

12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1265	99.1	1276	3	AAD00294 Mouse ooc
2	1245.4	97.5	1260	10	Adj63191 Mouse zyg
3	1234.2	96.6	1258	8	ABZ24589 Mouse ova
4	1234.2	96.6	1258	10	Adj63164 Mouse zyg
5	962.2	75.3	1280	10	Adj63200 Rat zygote
6	799.4	62.6	814	10	Adj63208 Zygote ar
7	799.4	62.6	873	8	ABZ24590 Mouse ova
8	799.4	62.6	873	10	Adj63203 Mouse zyg
9	799.4	62.6	873	10	Adj63174 Mouse zyg
10	753.8	59.0	4090	8	ABZ24591 Mouse ova
11	753.8	59.0	4090	10	Adj63175 Mouse zyg
12	628	49.2	3959	10	Adj63201 Rat zygote
13	352	27.6	1275	10	Adj63193 Human zyg
14	273.6	21.4	305	10	Adj63211 Zygote ar
15	205.8	16.1	1052	10	Adj63194 Frog zyg
16	194.8	15.3	963	10	Adj63198 Torafugu
17	194.6	15.2	1192	10	Adj63196 Zebrafish
18	134.6	10.5	2075	8	ABZ24592 Human ova
19	134.6	10.5	6002	10	Adj63204 Human zyg
20	134.6	10.5	7405	10	Adj63176 Human zyg
21	98.8	7.7	123	10	Adj63209 Zygote ar

22	72	5.6	105	10	ADJ63210	Zygote ar
23	58.8	4.6	114955	2	AAX53491	Human ade
24	57.8	4.5	2000	8	ADA71938	Rice gene
25	54.2	4.2	1032	8	ADA65887	Rice gene
26	54.2	4.2	1500	12	ADI42217	Plant tra
27	54.2	4.2	1500	12	ADO02720	Soybean o
28	54.2	4.2	4667	12	ADQ22939	Human sof
c 29 52 4.1 114955 2 AAX53491						
c 30 51.6 4.0 987 8 ABZ37562						
c 31 51.6 4.0 59816 8 ABZ37516						
c 32 51.6 4.0 59816 8 ABZ37515						
c 33 50.2 3.9 1673 12 ADI42224						
c 34 50.2 3.9 1673 12 ADO02727						
c 35 49.6 3.9 400 2 AAV44436						
c 36 49.6 3.9 400 2 AAV64545						
c 37 49.6 3.9 400 2 AAZ19134						
c 38 49.6 3.9 400 2 AAZ19346						
c 39 49.6 3.9 14800 6 ABL66291						
c 40 49.6 3.9 14800 12 ADP13447						
c 41 49.6 3.9 14835 6 AAS94858						
c 42 49 3.8 5235 12 ADM31027						
c 43 49 3.8 6792 12 ADM31025						
c 44 49 3.8 7175 2 AAQ84658						
c 45 49 3.8 7175 2 AAV42686						

ALIGNMENTS

RESULT 1

AAD00294

ID AAD00294 standard; cDNA; 1276 BP.

XX AC AAD00294;

XX DT 05-SEP-2000 (first entry)

XX DE Mouse oocyte-specific O1-180 cDNA clone.

XX KW Oocyte-specific; ovary; O1-180; mouse; gynaecological; treatment; screen;

XX KW cell proliferative disorder; cell degenerative disorder; contraceptive;

XX KW modulator; signalling pathway; human infertility; cancer; ovulation; ss.

XX OS Mus sp.

XX FH Key Location/Qualifiers

XX FT CDS 28..1113

XX FT /*tag= a

XX FT /product= "Mouse oocyte-specific protein, O1-180"

XX PN WO200024755-A1.

XX PD 04-MAY-2000.

XX PF 28-OCT-1999; 99WO-US025209.

XX PR 28-OCT-1998; 98US-0106020P.

XX PA (BAYU) BAYLOR COLLEGE MEDICINE.

XX PI Matzuk MM, Wang P;

XX DR WPI; 2000-350684/30.

XX P-PSDB; RAY70948.

XX O1-180, O1-184 and O1-236 polypeptides and nucleic acids encoding them, useful for evaluating potential contraceptives to block ovulation in a reversible manner.

XX PS Claim 2; Fig 1; 54pp; English.

XX CC The present sequence is the cDNA encoding the mouse oocyte-specific protein O1-180, expressed in the oocytes of primary (one-layer) preantral

XX Claim 1; SEQ ID NO 28; 219pp; English.
 PS The invention comprises the amino acid and coding sequences of ovary-
 CC specific (also known as zygote arrest 1 - Zar1) proteins and
 CC nucleoplamin (Npm2) proteins. The DNA and protein sequences of the
 CC invention are useful for inhibiting or enhancing early embryogenesis by
 CC disturbing the maternal genome. The DNA and protein sequences are also
 CC useful for screening mutations in components of those signalling pathways
 CC that are associated with some forms of human infertility or
 CC gynecological cancers. The present DNA sequence encodes a Zar1 protein.
 XX Sequence 1260 BP; 261 A; 390 C; 395 G; 214 T; 0 U; 0 Other;

Query Match 97.5%; Score 1245.4; DB 10; Length 1260;
 Best Local Similarity 99.8%; Pred. No. 5,1e-286;
 Matches 1257; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 3 GCGGGGAGGCGCGGACGCAACCATGTTCCCGCGAGCAGTTCACCCCTGCCGCA 62
 DB 3 GCGGGGAGGCGCGGACGCAACCATGTTCCCGCGAGCAGTTCACCCCTGCCGCA 62
 QY 63 TCCTTATCCGACGCGCACCAAGCCGGGATGGCTGGAGTTCGAGCAGGGGCTGCG 122
 DB 63 TCCTTATCCGACGCGCACCAAGCCGGGATGGCTGGAGTTCGAGCAGGGGCTGCG 122
 QY 123 ACCCGCGCCCTCTCTTCTCCCGGCTACAGACAGCTCATGGCGCGGAGTACGTGCA 182
 DB 123 ACCCGCGCCCTCTCTTCTCCCGGCTACAGACAGCTCATGGCGCGGAGTACGTGCA 182
 QY 183 CAGCCACGAGGGGACAGCTCATGGCCCTGCTGTCGCGATGGTCCCGTCCGTCAG 242
 DB 183 CAGCCACGAGGGGACAGCTCATGGCCCTGCTGTCGCGATGGTCCCGTCCGTCAG 242
 QY 243 CAGCGCTGACGCTGCGGTGAGGTGAACCCCGCGCGAGCAGCTCGGTGCTACT 302
 DB 243 CAGCGCTGACGCTGCGGTGAGGTGAACCCCGCGCGAGCAGCTCGGTGCTACT 302
 QY 303 CCGGGCGCGCAGCTGACGCTGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGG 362
 DB 303 CCGGGCGCGCAGCTGACGCTGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGG 362
 QY 363 CTGTCAACCCCGTGCGCAGCGCGCGCGGAGATCCCGCGATCCTGGCAGACCGTAGC 422
 DB 363 CTGTCAACCCCGTGCGCAGCGCGCGCGGAGATCCCGCGATCCTGGCAGACCGTAGC 422
 QY 423 CCGCTTCTCGTCCGTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 482
 DB 423 CCGCTTCTCGTCCGTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 482
 QY 483 GCAGACACCCAGAGGAGGAGGAGCGCGCATCTCTCGGAGCCCGGGAACCGGAGCC 542
 DB 483 GCAGACACCCAGAGGAGGAGGAGCGCGCATCTCTCGGAGCCCGGGAACCGGAGCC 542
 QY 543 GAGAGAGTGGCGCGAGGAAAGCGGTCCCCAGCCGCGAAGCGGAGGAGCGATGTTCA 602
 DB 543 GAGAGAGTGGCGCGAGGAAAGCGGTCCCCAGCCGCGAAGCGGAGGAGCGATGTTCA 602
 QY 603 GGCTCAGGCGAGGCGGGTGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTGGC 662
 DB 603 GGCTCAGGCGAGGCGGGTGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTGGC 662
 QY 663 GCGGATGAGTCTGAGCTGGAGCGAGGAGCAGTGTCTGCGCAGAGATGGCTCAGGA 722
 DB 663 GCGGATGAGTCTGAGCTGGAGCGAGGAGCAGTGTCTGCGCAGAGATGGCTCAGGA 722
 QY 723 CCGCGGTGATTCGATGCGCTTCGAGACGAGCCTCCCGCAAAAGCAGCGAGCAGCA 782
 DB 723 CCGCGGTGATTCGATGCGCTTCGAGACGAGCCTCCCGCAAAAGCAGCGAGCAGCA 782
 QY 783 GGAGCGCTCGTTCAGTCTTACGAGCAGAGTACGCTACTATCTACTGCAAGGACTG 842
 DB 783 GGAGCGCTCGTTCAGTCTTACGAGCAGAGTACGCTACTATCTACTGCAAGGACTG 842

QY 843 CAAATCCGTTGGGAGAGCGCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 902
 DB 843 CAAATCCGTTGGGAGAGCGCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 901
 QY 903 TCAACAGTCTCTGCGGAGTGTGTGAGAAATCCTACAACTTACAGAGTGGAGGACATCA 962
 DB 902 TCAACAGTCTCTGCGGAGTGTGTGAGAAATCCTACAACTTACAGAGTGGAGGACATCA 961
 QY 963 CCTGTCAAGTGTGTAAGAACTAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1022
 DB 962 CCTGTCAAGTGTGTAAGAACTAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1021
 QY 1023 CTAAACGCCCCATCGGCAAGACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1082
 DB 1022 CTAAACGCCCCATCGGCAAGACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1081
 QY 1083 ACAGACCTTCAGCTTCAATATACATCATTTAGTGTGTGTGTGTGTGTGTGTGTGTGT 1142
 DB 1082 ACAGACCTTCAGCTTCAATATACATCATTTAGTGTGTGTGTGTGTGTGTGTGTGTGT 1141
 QY 1143 GGGCTAATGGAATGGACAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1202
 DB 1142 GGGCTAATGGAATGGACAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1201
 QY 1203 TTCTATGACAGACAGT 1261
 DB 1202 TTCTATGACAGACAGT 1260

RESULT 3
 ABZ24589
 ID ABZ24589 standard; cDNA; 1258 BP.
 AC ABZ24589;
 DT 31-MAR-2003 (first entry)
 XX Mouse ovary-specific O1-180 cDNA.
 DE Ovary; O1-180; mouse; contraceptive; antiinfertility; cytostatic;
 KW gene therapy; gene; ss.
 OS Mus musculus.
 FH Key Location/Qualifiers
 FT CDS 26..1108
 FT /*tag= a
 FT /product= "Murine O1-180"
 FT /transl_except= (pos:182..184,aa:Ser)
 FT /transl_except= (pos:1004..1007,aa:Phe)
 XX WO200288314-A2.
 XX 07-NOV-2002.
 XX PD 26-APR-2002; 2002WO-US013245.
 XX PF 27-APR-2001; 2001US-00844864.
 XX PR (BAYU) BAYLOR COLLEGE MEDICINE.
 XX PA (AMHP) WYETH.
 XX PI Matzuk MM, Wang P, Bai Y, Wu X;
 XX WPI; 2003-167110/16.
 DR P-PSDB; ABP58235.
 XX New ovary-specific-genes comprising O1-180 or O1-236, useful for
 PT decreasing conception or enhancing fertility, or for the preparation of a
 PT composition for treating e.g. cancer.
 XX Disclosure; Page 106-107; 141pp; English.

The present sequence is that of murine ovary-specific OI-180 cDNA obtained from a mouse ovary cDNA library. OI-180 clones were initially identified in a subtractive hybridisation screen using ovaries from Gdf9 knockout and wild-type mice. The OI-180 gene (see AB224590) on chromosome 5 is also provided. Loss of OI-180 results in female infertility and subfertility. The invention provides ovary-specific and oocyte-specific murine and human OI-180, OI-184 and OI-236 polynucleotides and polypeptides. These genes and their protein products appear to relate to various cell proliferative or degenerative disorders, especially those involving ovarian tumours, such as germ line tumours and granulosa cell tumours, or infertility, such as premature ovarian failure. The invention provides a method for detection of a cell proliferative or degenerative disorder of the ovary, which is associated with the expression of OI-180, OI-184 or OI-236. It also provides a method for treating such disorders by using an agent which suppresses or enhances the respective activities of OI-180, OI-184 or OI-236, and a method of screening for compounds that interact and/or modulate the expression or activity of the ovary-specific genes. These compounds are possible contraceptive or fertility enhancing agents. The modulator is preferably a polypeptide, small molecule or polynucleotide sequence.

IX
20 sequence 1258 bp. 261 A: 392 C: 392 G: 213 T: 0 U: 0 Other;

Query Match	96.6%;	Score 1234.2;	DB 8;	Length 1358;
Best Local Similarity	99.3%;	Pred. No. 2.4e-283;		
Matches 1250;	Conservative	0;	Mismatches 8;	Indels 1;
				Gaps 1;

3	GGCGGCGAGGCGCGGAGCGCACCCATGTTCCGCGAGCAGCTTCCACCCCTGCGCCGCA	62
1	GGCGGGGCGAGCGCGGGGAGCGACCCATGTTCCGCGGAGCAGTTCCACCCCTGCGCCGCA	60
63	TCCTTTATCCGACAGGCCACAAAGCCGGGAGATGGCTGGAGGTTTCGAGCAGCAGGGCTGCGC	122
61	TCCTTTATCCGACAGGCCACAAAGCCGGGAGATGGCTGGAGGTTTCGAGCAGCAGGGGTCGCG	120
123	ACCGCGCCCCCTCTCTTCTCCCGCTACAGACAGCTCATGCGCGCGGAGTAGTCTCGA	182
121	ACCGCGCCCCCTCTCTTCTCCCGCTACAGACAGCTCATGCGCGCGGAGTAGTCTCGA	180
183	CAGCCACAGCGGGGACAGCTCATGCGCCCTGCTGTGCGGATGGCTCCCGTTCGGTCAG	242
181	CCGCGCACAGCGGGGACAGCTCATGCGCCCTGCTGTGCGGATGGGTCCCGTTCGGTCAG	240
243	CAGCCGTGACGCTGCGGTGCAGGTGAACCCGCGCGCGACGCTCGGTTCAGTGTTCAC	302
241	CAGCCGTGACGCTGCGGTGCAGGTGAACCCGCGCGCGACGCTCGGTTCAGTGTTCAC	300
303	CGGCGCGGACACGCTGCAGCTGCAGGGTGCAGCAGCAGCCGACGCGCCGATCGGGTTC	362
301	CGGCGCGGACACGCTGCAGGCTGCAGGGTGCAGCAGCAGCCGACGCGCCGATCGGGTTC	360
363	CTGTCAACCCCGTGGCCACGCGCGCGCGGGAGATCCCGCGATCTTGCGACACCGTTCAG	422
361	CTGTCAACCCCGTGGCCACGCGCGCGCGGGAGATCCCGCGATCTTGCGACACCGTTCAG	420
423	CCGCTTCTCGTCCGTGACCTTCTGTGCGCTCTCTCTCACTGGAGGTTTCGGGAGGCGAG	482
421	CCGCTTCTCGTCCGTGACCTTCTGTGCGCTCTCTCTCACTGGAGGTTTCGGGAGGCGAG	480
483	GCAGACACCCACGAAGGGAGAGGGGAGCCCGGATCTCTGGGAGACCCGGGAACCGGAGCC	542
481	GCAGACACCCACGAAGGGAGAGGGGAGCCCGGATCTCTGGGAGACCCGGGAACCGGAGCC	540
543	GAGAGAGTGGCGCGAGGAAGCGGTCCGCCAGCCGCGAAGCGAGAGGGCGGATTTCA	602
541	GAGAGAGTGGCGCGAGGAAGCGGTCCGCCAGCCGCGAAGCGAGAGGGCGGATTTCA	600
603	GGCTCAGGGGACGCGCGGTGGAGCGAGCAGCAGCAGCCGAGGACCGGACAGTCTGGC	662
601	GGCTCAGGGGACGCGCGGTGGAGCGAGCAGCAGCAGCCGAGGACCGGACAGTCTGGC	660
663	GGCGATGCACTGTGAGCTGGGAGCGAGGAGGCCATGTCTCTGCCGACAGATGGCTCAGGA	722

DR WPI; 2003-865579/80.
 XX P-PSDB; ADJ63165.
 PT New isolated ovary-specific genes e.g. OI-180, OI-184 or OI-236, useful
 in inhibiting or enhancing early embryogenesis or treating cell
 PT proliferative or degenerative disorders which are mediated by OI-180, OI-
 PT 184 or OI-236 proteins.
 XX
 PS Disclosure; SEQ ID NO 1; 219pp; English.
 XX
 CC The invention comprises the amino acid and coding sequences of ovary-
 CC specific (also known as zygote arrest 1 - Zar1) proteins and
 CC nucleoplasmin (Npm2) proteins. The DNA and protein sequences of the
 CC invention are useful for inhibiting or enhancing early embryogenesis by
 CC disturbing the maternal genome. The DNA and protein sequences are also
 CC useful for screening mutations in components of those signalling pathways
 CC that are associated with some forms of human infertility or
 CC gynaecological cancers. The present DNA sequence encodes a Zar1 protein.
 XX
 SQ Sequence 1258 BP; 261 A; 392 C; 392 G; 213 T; 0 U; 0 Other;
 Query Match 96.6%; Score 1234.2; DB 10; Length 1258;
 Best Local Similarity 99.3%; Pred. No. 2.4e-283;
 Matches 1250; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
 QY 3 GCGGGCGAGGCGCGGACGACACCATCTTCCCGGAGACAGCTTCCACCCCTGCCCCGCA 62
 DB 1 GCGGGCGAGGCGCGGAGCGACCATCTTCCCGGAGACAGCTTCCACCCCTGCCCCGCA 60
 QY 63 TCCTTATCCGAGGCGACCAAGCGGGGATGGCTGGAGGTTGCGAGCCAGGGGTCGCG 122
 DB 61 TCCTTATCCGAGGCGACCAAGCGGGGATGGCTGGAGGTTGCGAGCCAGGGGTCGCG 120
 QY 123 ACCGCGCCCT 182
 DB 121 ACCGCGCCCT 180
 QY 183 CAGCCACCGAGGCGACAGCTCATGGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 242
 DB 181 CCGCCACCGAGGCGACAGCTCATGGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
 QY 243 CAGCGGTGACGTGCGGTGCGAGTGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 302
 DB 241 CAGCGGTGACGTGCGGTGCGAGTGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
 QY 303 CCGGCG 362
 DB 301 CCGGCG 360
 QY 363 CTGTCAACCCCGTGGCCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 422
 DB 361 CTGTCAACCCCGTGGCCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
 QY 423 CCGGTTCTGTCGCGTCACTTCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 482
 DB 421 CCGGTTCTGTCGCGTCACTTCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
 QY 483 CGAGACCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 542
 DB 481 CGAGACCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
 QY 543 GAGAGAGTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 602
 DB 541 GAGAGAGTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
 QY 603 GGCTGACGGGAGGCGCGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 662
 DB 601 GGCTGACGGGAGGCGCGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
 QY 663 GGCGATGAGTGTGAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 722
 DB 661 GGCGATGAGTGTGAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720

QY 723 CCCGGTGATTCGGATGCCCTCGAGACCAAGCCCTCCCGCAAAAGCAGGAGGAGCA 782
 DB 721 CCCGGTGATTCGGATGCCCTCGAGACCAAGCCCTCCCGCAAAAGCAGGAGGAGCA 780
 QY 783 GGAGCGCTGGTTTCCAGTTCTTAGAGCAAGAGTACGGCTACTATCACTGCAAGGACTG 842
 DB 781 GGAGCGCTGGTTTCCAGTTCTTAGAGCAAGAGTACGGCTACTATCACTGCAAGGACTG 840
 QY 843 CAAAATCCGGTGGGAGAGCGCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 902
 DB 841 CAAAATCCGGTGGGAGAGCGCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 899
 QY 903 TCAAAACAGTTCTCGCGGAGTGTGTGAGAAATCCTTACAACCTTTACAGAGTGGAGGACATCA 962
 DB 900 TCAAAACAGTTCTCGCGGAGTGTGTGAGAAATCCTTACAACCTTTACAGAGTGGAGGACATCA 959
 QY 963 CCTGTCAAAGTTGTAAGAACTAGATGTGCTGCCAGTGCAGATTTGCCACGTTGGAGC 1022
 DB 960 CCTGTCAAAGTTGTAAGAACTAGATGTGCTGCCAGTGCAGATTTGCCACGTTGGAGC 1019
 QY 1023 CTAAACGCCCCATCGCAAGACTTGTGTGGAGATGCAAGGACAAACCGCTGTCTCTCG 1082
 DB 1020 CTAAACGCCCCATCGCAAGACTTGTGTGGAGATGCAAGGACAAACCGCTGTCTCTCG 1079
 QY 1083 ACAGCACCTTCAGCTTCAAATACATCAATTTAGTGAGAGTCAAAAACGTTTCTGCTAGATG 1142
 DB 1080 ACAGCACCTTCAGCTTCAAATACATCAATTTAGTGAGAGTCAAAAACGTTTCTGCTAGATG 1139
 QY 1143 GGGCTAATGGAATGGACAGAGTGTCTCCCTCTTTCACCTCTTCCCTTTCCAAATTC 1202
 DB 1140 GGGCTAATGGAATGGACAGAGTGTCTCCCTCTTTCACCTCTTCCCTTTCCAAATTC 1199
 QY 1203 TTCTATGACAGACAGT 1261
 DB 1200 TTCTATGACAGAGT 1258
 RESULT 5
 ID ADJ63200 standard; DNA; 1280 BP.
 XX
 AC ADJ63200;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Rat zygote arrest 1 (Zar1) DNA sequence.
 XX
 KW ovary-specific protein; zygote arrest 1; Zar1; nucleoplasmin; Npm2;
 KW embryogenesis; mutation screening; infertility; gynaecological cancer;
 KW ds; rat.
 XX
 OS Rattus norvegicus.
 XX
 PN WO2003091400-A2.
 XX
 PD 06-NOV-2003.
 XX
 PF 23-APR-2003; 2003WO-US012720.
 XX
 PR 26-APR-2002; 2002WO-US013245.
 PR 17-SEP-2002; 2002US-0411262P.
 PR 17-DEC-2002; 2002US-0434165P.
 PR 13-JAN-2003; 2003US-0439781P.
 PR 23-JAN-2003; 2003US-0442164P.
 XX
 (BAYU) BAYLOR COLLEGE MEDICINE.
 PA (AMHP) WYETH.
 XX
 PI Matzuk MM, Wang P, Wu X, Bai Y;
 XX
 DR WPI; 2003-865579/80.
 XX
 PT New isolated ovary-specific genes e.g. OI-180, OI-184 or OI-236, useful

PT in inhibiting or enhancing early embryogenesis or treating cell
PT proliferative or degenerative disorders which are mediated by O1-180, O1-
PT 184 or O1-236 proteins.

Claim 1: SEO ID NO 37: 219pp; English.

The invention comprises the amino acid and coding sequences of ovary-specific (also known as zygote arrest 1 - Zar1) proteins and of nucleoplasmin (Npm2) proteins. The DNA and protein sequences of the invention are useful for inhibiting or enhancing early embryogenesis by disturbing the maternal genome. The DNA and protein sequences are also useful for screening mutations in components of those signalling pathways that are associated with some forms of human infertility or gynaecological cancers. The present nucleic acid represents a Zar1 DNA sequence.

sequence 1280 BP: 264 A; 408 C; 396 G; 212 T; 0 U; 0 Other;

Query Match	75.3%	Score 962.2;	DB 10;	Length 1280;
Best Local Similarity	87.7%	Pred. No. 1.1e-218;		
Matches 1110: Conservative	0;	Mismatches 143;	Indels 12;	Gaps 5;

QY	3	GGCGGCGAGCGCGGAGACGACCCATGTTCCCGGAGACACGTTCCACCCCTGCCCGCA	62
DB	3	GGCGGCGAGCGCGGAGACGACCCATGTTCCCGGAGACACGTTCCACCCCTGCCCGCA	62
QY	63	TCCTTATCCGC- --AGGCCACAAAGCCGGGATGGCTGGAGGTTTCGAGCAGCGGGCTG	119
DB	63	TCCTTATCCCGCCACGGCAGCAAGCCGGGATGGCTGGAGGTTTGAGCAGCGGGCTG	122
QY	120	CCGACCCGCGCCGCTCTTCCTCCCGGCTACAGACAGCTCATGGCCGGGAGTAGT	179
DB	123	CAGGCCGAGCCGCTCTTCCTCCCGGCTACAGACAGCTCATGGCCGGGAGTAGT	182
QY	180	CGACAGCACAGCGGGGACAGCTCATGCCCTGCTGTCGGGATGGGTCCCGGTCGGT	239
DB	183	TGACAGCTATCAGCGAGCGAGCTCATGGCTTGTCTGTCGAAATGGTTCGCCGCGGT	242
QY	240	CAGCAGCGGTGACGTGCGGTGACAGTGAACCCGCGCGCGACAGCCTCGGTGCAGTTC	299
DB	243	CAGCAGCCGACGCTGCGGTGACAGTGAACCCGCGCGCGAGTCCCTCGGTGCAGTTC	302
QY	300	ACTCGGGCGCGCAGCTGCAGCCTGCAGGCTGCCGAGCCAGCCCGCCGATCCGG	359
DB	303	GCTCGGGCGCGCACATGACGCTTGACGCGCGCGAGCCAGCCCGCGCTCG	362
QY	360	TTCTCTGTCAAACCCCGTGCCACGCGCGCGCGGGAGATCCCGCGATCCTGGCAGCCGT	419
DB	363	TTCTCTGCCAAACCCCGTGCCAGCCCGCGCGGGAGACCCCGCGATCCTGGCGACCGT	422
QY	420	AGCCCGGTCTCGTCCGTGACCTTCTGTGGCTCTCTCTCTCACGTGAGTTGCGGAGG	479
DB	423	CGCCCTGTACTCGCCGCTGACCTTCGTGGCTCTCTCTCTCGTGGAGTTGCGGGA	482
QY	480	CAGGCAGACACCAAGAGGAGGGGAGCCCGCATCCTCGGGAACCGGAACCGGA	539
DB	483	CAGGCAGACGCCACGAAGGAGGGGAGACCGGCACCCAGGGGACCCGGGAACCGGA	542
QY	540	GCCGAGAGGTGGCCGACAGGAACGCTTCCCGCAGCCGCGAGCGAGGGGCGATCT	599
DB	543	GCCGGAGAGGTGGCAGTATGAAGACATGTCCTCCCGCGAGCGAGCGAGGGGCGCT	602
QY	600	TGAGGCTGCAGGCGAGCCCGGTGGGAGCAGCCACCAACCGGAGCCGGAACAGTGT	659
DB	603	CCAGGCTGAAGGCGAGGATG- --GGCAGGAGCAGCCACCCGGGAGGACCCCGACAGTGT	659
QY	660	GGCGGCGATGCTGAGCTGGGAGCGAGGACCATGCTCTGCGCAGAGATGCTCA	719
DB	660	GGCGGCGATGCTGAGCTGAGCCCGGAGTAGGAGGCCCTCTCTGCTGTCGAGATGCTCA	719
QY	720	GGACCCCGGTGATTGGATGCCCTCGAGACCAAGGCTCCCGCAAGCAGCAGGAGGA	779
DB	720	GGACCCCGGTGATTGGATGCCCTCGAGACCAAGGCTCCCGCAAGCAGCAGGAGGA	779

PT New isolated ovary-specific genes e.g. O1-180, O1-184 or O1-236, useful
PT in inhibiting or enhancing early embryogenesis or treating cell
PT proliferative or degenerative disorders which are mediated by O1-180, O1-
PT 184 or O1-236 proteins.

PS Disclosure; Fig 19; 219pp; English.

XX The invention comprises the amino acid and coding sequences of ovary-
CC specific (also known as zygote arrest 1 - Zar1) proteins and
CC nucleoplamin (Npm2) proteins. The DNA and protein sequences of the
CC invention are useful for inhibiting or enhancing early embryogenesis by
CC disturbing the maternal genome. The DNA and protein sequences are also
CC useful for screening mutations in components of those signalling pathways
CC that are associated with some forms of human infertility or
CC gynaecological cancers. The present DNA sequence represents exon 1 from a
CC Zar1 gene.

SQ Sequence 814 BP; 135 A; 285 C; 290 G; 104 T; 0 U; 0 Other;

Query Match 62.6%; Score 799.4; DB 10; Length 814;
Best Local Similarity 99.9%; Pred. No. 5e-180;
Matches 800; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	3	GGCGGCGAGCGCGGAGCGACACCATGTTCCCGGCGAGCAGCTTCCACCCCTGCGCGCA	62
Db	1	GGCGGCGAGCGCGGAGCGACACCATGTTCCCGGCGAGCAGCTTCCACCCCTGCGCGCA	60
Qy	63	TCCTTTATCCGCGAGGCCACCAAGCGGGGATGCGTGGAGGTTCCGAGCAGGGCTGCGG	122
Db	61	TCCTTTATCCGCGAGGCCACCAAGCGGGGATGCGTGGAGGTTCCGAGCAGGGGCTGCGG	120
Qy	123	ACCGCGCGCCCTCTCTTCTCTCCCGCTACAGACAGCTCATGGCGCGGAGTACGTGCA	182
Db	121	ACCGCGCGCCCTCTCTTCTCTCCCGCTACAGACAGCTCATGGCGCGGAGTACGTGCA	180
Qy	183	CAGCCACACAGCGGGCAGAGCTATGGCCCTCTGTGCGGATGGTCCCGGTGGGTGAG	242
Db	181	CAGCCACACAGCGGGCAGAGCTATGGCCCTCTGTGCGGATGGTCCCGGTGGGTGAG	240
Qy	243	CAGCGGTGACCTGCGGTGCAAGTGAACCGCGCGCGGACCGCTCGGTGCAAGTGTACT	302
Db	241	CAGCGGTGACCTGCGGTGCAAGTGAACCGCGCGCGGACCGCTCGGTGCAAGTGTACT	300
Qy	303	CGGCGCGCGCACGCTGACAGCTGACAGGTGCGGAGCCAGCCCGACCGCCGATCGGPTC	362
Db	301	CGGCGCGCGCACGCTGACAGCTGACAGGTGCGGAGCCAGCCCGACCGCCGATCGGPTC	360
Qy	363	CTGTCAACCCCGTGCGCCACGCGCGCGCGGAGATCCCGGATCTCTGGCAGACCGTAGC	422
Db	361	CTGTCAACCCCGTGCGCCACGCGCGCGCGGAGATCCCGGATCTCTGGCAGACCGTAGC	420
Qy	423	CCCGTTCTCGTCCGTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	482
Db	421	CCCGTTCTCGTCCGTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	480
Qy	483	CGACACACCCACGAGGAGGAGGAGGAGCCCGGATCTCTCGGAGCCCGGAGACCGAGCC	542
Db	481	CGACACACCCACGAGGAGGAGGAGGAGCCCGGATCTCTCGGAGCCCGGAGACCGAGCC	540
Qy	543	GAGAGAGTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	602
Db	541	GAGAGAGTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	600
Qy	603	GGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	662
Db	601	GGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	660
Qy	663	GCGCATGAGTGTGAGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	722
Db	661	GCGCATGAGTGTGAGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	720
Qy	723	CCCGCGTATTCGATGCGGCTTCGAGACCAAGCCCTCCCGCAAGACCGAGCAGGACAA	782

Db	721	CCCGCGTATTCGATGCGGCTTCGAGACCAAGCCCTCCCGCAAGACCGAGCAGGACAA	780
Qy	783	GGAGCGCTGGTTTCCAGTT	803
Db	781	GGAGCGCTGGTTTCCAGTT	801

RESULT 7

ID	ABZ24590	standard; DNA; 6873 BP.
XX	ABZ24590;	
AC	ABZ24590;	
DT	31-MAR-2003	(first entry)
DE	Mouse ovary-specific O1-180 gene.	
XX	Ovary; O1-180; mouse; contraceptive; antiinfertility; cytostatic;	
KW	gene therapy; chromosome 5; gene; ds.	
XX	Mus musculus.	
OS		
FH	Key	Location/Qualifiers
FT	exon	2536..3311
FT		/*tag= b
FT		/number= 1
FT	CDS	2538..6353
FT		/*tag= a
FT		/product= "Murine O1-180"
FT		/note= "contains introns"
FT	intron	3312..5083
FT		/*tag= c
FT		/number= 1
FT	exon	5084..5177
FT		/*tag= d
FT		/number= 2
FT	intron	5178..5254
FT		/*tag= e
FT		/number= 2
FT	exon	5255..5324
FT		/*tag= f
FT		/number= 3
FT	intron	5325..6206
FT		/*tag= g
FT		/number= 3
FT	exon	6207..6352
FT		/*tag= h
FT		/number= 4
XX	WO200288314-A2.	
PD	07-NOV-2002.	
XX		
PF	26-APR-2002; 2002WO-US013245.	
XX		
PR	27-APR-2001; 2001US-00844864.	
XX		
PA	(BAYU) BAYLOR COLLEGE MEDICINE.	
PA	(AMHP) WYETH.	
XX		
PI	Matzuk MM, Wang P, Bai Y, Wu X;	
XX		
DR	WPI; 2003-167110/16.	
DR	P-PSDB; ABP58235.	
XX		
PT	New ovary-specific-genes comprising O1-180 or O1-236, useful for	
PT	decreasing conception or enhancing fertility, or for the preparation of a	
PT	composition for treating e.g. cancer.	
XX		
PS	Claim 1; Page 130-134; 141pp; English.	
XX		
CC	The present sequence is that of the murine ovary-specific O1-180 gene on	
CC	chromosome 5. O1-180 clones were initially identified in a cDNA	

Db 601 GGCTGAGGCGGCGGGTGGAGCAGCAGCCACCGAGGACCGGAACAGTGTGC 660
 Qy 663 GCGGATGAGTCTGAGCCTGGGAGCGAGGAGCATGTCTGCGCGAGAGATGGCTCAGA 722
 Db 661 GCGGATGAGTCTGAGCCTGGGAGCGAGGAGCATGTCTGCGCGAGAGATGGCTCAGA 720
 Qy 723 CCCCGGTGATTCGGATGCCCTTCGAGACCCAGCCCTCCCGCAAGCAAGCAGAGCAAA 782
 Db 721 CCCCGGTGATTCGGATGCC-----CCTCCCGCAAGCAAGCAGAGCAAA 767
 Qy 783 GGAGCGCTGCGTTTCCAGTT 803
 Db 768 GGAGCTCCTGCGTTTCCAGGT 788

RESULT 11
 ADJ63175
 ID ADJ63175 standard; DNA; 4090 BP.
 XX AC ADJ63175;
 XX DT 06-MAY-2004 (first entry)
 DE Mouse zygote arrest 1 (Zarl) DNA sequence #2.
 KW ovary-specific protein; zygote arrest 1; Zarl; nucleoplasmin; Npm2;
 KW embryogenesis; mutation screening; infertility; gynaecological cancer;
 KW ds; mouse; murine.
 XX Mus musculus.
 XX PN WO2003091400-A2.
 XX PD 06-NOV-2003.
 XX PF 23-APR-2003; 2003WO-US012720.
 XX PR 26-APR-2002; 2002WO-US013245.
 XX PR 17-SEP-2002; 2002US-0411262P.
 XX PR 17-DEC-2002; 2002US-0434165P.
 XX PR 13-JAN-2003; 2003US-0439781P.
 XX PR 23-JAN-2003; 2003US-0442164P.
 XX (BAYU) BAYLOR COLLEGE MEDICINE.
 XX (AMHP) WYETH.
 XX PA Matzuk MM, Wang P, Wu X, Bai Y;
 XX PI WPI; 2003-865579/80.
 XX DR New isolated ovary-specific genes e.g. OI-180, OI-184 or OI-236, useful
 XX PT in inhibiting or enhancing early embryogenesis or treating cell
 XX PT proliferative or degenerative disorders which are mediated by OI-180, OI-
 XX PT 184 or OI-236 proteins.
 XX PS Claim 1; SEQ ID NO 12; 219pp; English.

CC The invention comprises the amino acid and coding sequences of ovary-
 CC specific (also known as zygote arrest 1 - Zarl) proteins and
 CC nucleoplasmin (Npm2) proteins. The DNA and protein sequences of the
 CC invention are useful for inhibiting or enhancing early embryogenesis by
 CC disturbing the maternal genome. The DNA and protein sequences are also
 CC useful for screening mutations in components of those signalling pathways
 CC that are associated with some forms of human infertility or
 CC gynaecological cancers. The present nucleic acid represents a Zarl DNA
 CC sequence.
 XX SQ Sequence 4090 BP; 274 A; 408 C; 413 G; 239 T; 0 U; 2756 Other;
 Query Match 59.0%; Score 753.8; DB 10; Length 4090;
 Best Local Similarity 97.5%; Pred. No. 5.9e-169;
 Matches 781; Conservative 0; Mismatches 7; Indels 13; Gaps 1;

Qy 3 GCGGCGGAGCGCGGAGCGCACCATGTTCCCGGCGAGCAGTTCACCCCTGCCCGCA 62
 Db 1 GCGGCGGAGCGCGGAGCGCACCATGTTCCCGGCGAGCAGTTCACCCCTGCCCGCA 60
 Qy 63 TCCTTATCCGAGGCGCACAAAGCGGGATGGCTGGAGTTTCGAGACGAGGGTGGCG 122
 Db 61 TCCTTATCCGAGGCGCACAAAGCGGGATGGCTGGAGTTTCGAGACGAGGGTGGCG 120
 Qy 123 ACCGCGGCGCGCTTCTCTCCCGGCTACAGACAGCTCATGGCCGCGGAGTAGCTCGA 182
 Db 121 ACCGCGGCGCGCTTCTCTCCCGGCTACAGACAGCTCATGGCCGCGGAGTAGCTCGA 180
 Qy 183 CAGCCACAGCGGCGCACAGCTCATGGCCGCTGCTGCGCGGATGGGTCCCGTCCAG 242
 Db 181 CAGCCACAGCGGCGCACAGCTCATGGCCGCTGCTGCGCGGATGGGTCCCGTCCAG 240
 Qy 243 CAGCGTGACCTGCGGTGCGAGTGAACCCGCGCGGACCGCTCGGTGCGAGTGTCACT 302
 Db 241 CAGCGTGACCTGCGGTGCGAGTGAACCCGCGCGGACCGCTCGGTGCGAGTGTCACT 300
 Qy 303 CCGGCGGCGCGAGCTGCGAGCTGCGAGCTGCGAGCTGCGAGCTGCGAGCTGCGAGCT 362
 Db 301 CCGGCGGCGCGAGCTGCGAGCTGCGAGCTGCGAGCTGCGAGCTGCGAGCTGCGAGCT 360
 Qy 363 CTGTCAACCCCGTGGCGCACCGCGCGGAGATCCCGCGATCCTGGCAGACCGTAGC 422
 Db 361 CTGTCAACCCCGTGGCGCACCGCGCGGAGATCCCGCGATCCTGGCAGACCGTAGC 420
 Qy 423 CCGGTTCTGCTCGGTGAGCTTCTGTGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 482
 Db 421 CCGGTTCTGCTCGGTGAGCTTCTGTGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
 Qy 483 GCAGACACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 542
 Db 481 GCAGACACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
 Qy 543 GAGAGAGTGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 602
 Db 541 GAGAGAGTGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
 Qy 603 GGTGCGAGGCGAGCGCGGTGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 662
 Db 601 GGTGCGAGGCGAGCGCGGTGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 660
 Qy 663 GCGATGCACTGAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 722
 Db 661 GCGATGCACTGAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
 Qy 723 CCCCGGTGATTCGGATGCCCTTCGAGACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 782
 Db 721 CCCCGGTGATTCGGATGCC-----CCTCCCGCAAGCAAGCAGAGCAAA 767
 Qy 783 GGAGCGCTGCGTTTCCAGTT 803
 Db 768 GGAGCTCCTGCGTTTCCAGGT 788

RESULT 12
 ADJ63201
 ID ADJ63201 standard; DNA; 3959 BP.
 XX AC ADJ63201;
 XX DT 06-MAY-2004 (first entry)
 DE Rat zygote arrest 1 (Zarl) protein coding sequence #1.
 KW ovary-specific protein; zygote arrest 1; Zarl; nucleoplasmin; Npm2;
 KW embryogenesis; mutation screening; infertility; gynaecological cancer;
 XX gene; ds; rat.
 OS Rattus norvegicus.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2005, 16:10:41 ; Search time 128 Seconds
(without alignments)
7091.230 Million cell updates/sec

Title: US-09-830-810A-1
Perfect score: 1277
Sequence: 1 aagcgggcgagcgcgaggga.....acaaaaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	4.2	5467	1	US-07-745-206A-12
2	54	4.2	5467	2	US-08-311-363-12
3	53.2	4.2	7218	1	US-08-232-463-14
4	50.8	4.0	6232	3	US-08-456-200B-11
5	49.6	3.9	400	3	US-09-056-556-179
6	49.6	3.9	400	4	US-09-072-596-174
7	49.6	3.9	400	4	US-09-072-967-179
8	49	3.8	7175	1	US-08-455-543A-8
9	49	3.8	7175	2	US-08-193-078B-8
10	49	3.8	7175	2	US-08-223-305C-8
11	49	3.8	7175	2	US-08-149-097D-8
12	49	3.8	7175	3	US-08-949-386-8
13	49	3.8	7175	3	US-08-450-582-8
14	49	3.8	7175	3	US-08-984-709A-8
15	49	3.8	7175	3	US-08-450-272-8
16	49	3.8	7175	4	US-08-450-273-8
17	49	3.8	7177	3	US-09-268-163-7
18	49	3.8	7266	3	US-08-713-118-1
19	49	3.8	7266	3	US-09-452-007-1
20	49	3.8	7362	1	US-08-455-543A-7
21	49	3.8	7362	2	US-08-193-078B-7
22	49	3.8	7362	2	US-08-223-305C-7
23	49	3.8	7362	2	US-08-149-097D-7
24	49	3.8	7362	3	US-08-949-386-7
25	49	3.8	7362	3	US-08-450-582-7
26	49	3.8	7362	3	US-08-984-709A-7
27	49	3.8	7362	3	US-08-450-272-7

28	49	3.8	7362	4	US-08-450-273-7
29	49	3.8	7364	3	US-09-268-163-5
30	49	3.8	7376	3	US-09-268-163-3
31	48	3.8	4403765	3	US-09-103-840A-2
C 32	46.6	3.6	745	1	US-08-036-555B-163
C 33	46.6	3.6	745	1	US-08-469-569-163
C 34	46.6	3.6	745	1	US-08-249-322A-163
C 35	46.6	3.6	745	1	US-08-469-526A-163
C 36	46.6	3.6	745	2	US-08-734-591A-163
C 37	46.6	3.6	745	2	US-08-469-660-163
C 38	46.6	3.6	745	3	US-08-341-018-51
C 39	46.6	3.6	745	3	US-08-470-335-163
C 40	46.6	3.6	745	3	US-08-735-021-163
C 41	46.6	3.6	745	3	US-08-734-664A-163
C 42	46.6	3.6	745	3	US-08-470-339-163
C 43	46.6	3.6	745	4	US-08-467-602-163
C 44	46.6	3.6	745	4	US-08-411-295F-44
C 45	46.6	3.6	745	5	PCT-US94-05083C-159

ALIGNMENTS

RESULT 1
US-07-745-206A-12
; Sequence 12, Application US/07745206A
; Patent No. 5429921
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: McCue, Ann
; APPLICANT: Feldman, Daniel
; TITLE OF INVENTION: Human Calcium Channel Compositions and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fitch, Even, Tabin & Flannery
; STREET: 135 S. LaSalle
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07745,206A
; FILING DATE: 19910815
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Feder, Scott B
; REFERENCE/DOCKET NUMBER: 51504
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-372-7842
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5467 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(144..3164, 3168..3245, 3249..3386, 3390
; LOCATION: ..3392, 3396..3488, 3495..3539, 3543..3581, 3585
; LOCATION: ..3587, 3591..3626, 3630..3689, 3693..3737, 3744
; LOCATION: ..3746, 3750..4823, 4827..4841, 4845..5006, 5010
; LOCATION: ..5096, 5100..5306, 5310..5366, 5370..5465)
US-07-745-206A-12


```
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;
QY 250 GACGCTGCGGTGACGTGAACCCGCGCGCGCGCGCGCTCGGTGCAAGTGTCTACTCGGGCGC 309
Db 2751 GAGGCCCGAAGCGGAGAGCGGGAGCCCGGTGCGCGGAGGAGCGCGCGCGCGCAC 2810
QY 310 CGCAGCTGACCTGTCAGGTGTCGAGCAGCCCGAGCGCGCGATCGGTTCCTGTCAA 369
Db 2811 CGCA-GCCACAAAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2867
QY 370 CCCCGTGGCGACGCGCGCGCGCGGAGATCCCGCGCATCTGCGCACAGCTAGCCCGCTC 429
Db 2868 GCGCCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2927
QY 430 TCGTCGCTGACCTTCTGTGGCTCTCTCTCTCACTGAGAGTTCGCGGAGGCGAGCACA 489
Db 2928 GAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2987
QY 490 CCCAGAGGAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 549
Db 2988 GCGCCAAAGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3047
QY 550 GTGGCGCGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAT 609
Db 3048 GCGGAGAGCGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3107
QY 610 GCGGAGCGCGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAT 668
Db 3108 CACGAGGCTGTGAGAGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAT 3166

RESULT 9
US-08-193-078B-8
; Sequence 8, Application US/08193078B
; Patent No. 5846757
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWN, MARTIN, HALLER & MCCLAIN
; STREET: 1660 UNION STREET
; CITY: SAN DIEGO
; STATE: CA
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/193,078B
; FILING DATE: 07-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/868,354
; FILING DATE: 10-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-53607
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
```

```
TELEFAX: 619-238-0062
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7175 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 144..6857
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..143
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 6855..7175
US-08-193-078B-8

Query Match 3.8%; Score 49; DB 2; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0091;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;
QY 250 GACGCTGCGGTGACGTGAACCCGCGCGCGCGCGCGCTCGGTGCAAGTGTCTACTCGGGCGC 309
Db 2751 GAGGCCCGAAGCGGAGAGCGGGAGCCCGGTGCGCGGAGGAGCGCGCGCGCGCAC 2810
QY 310 CGCAGCTGACCTGTCAGGTGTCGAGCAGCCCGAGCGCGCGATCGGTTCCTGTCAA 369
Db 2811 CGCA-GCCACAAAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2867
QY 370 CCCGTGGCGACGCGCGCGCGCGGAGATCCCGCGCATCTGCGCACAGCTAGCCCGCTC 429
Db 2868 GCGCCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2927
QY 430 TCGTCGCTGACCTTCTGTGGCTCTCTCTCTCACTGAGAGTTCGCGGAGGCGAGCACA 489
Db 2928 GAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2987
QY 490 CCCAGAGGAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 549
Db 2988 GCGCCAAAGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3047
QY 550 GTGGCGCGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAT 609
Db 3048 GCGGAGAGCGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3107
QY 610 GCGGAGCGCGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAT 668
Db 3108 CACGAGGCTGTGAGAGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAT 3166

RESULT 10
US-08-223-305C-8
; Sequence 8, Application US/08223305C
; Patent No. 5851824
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/223,305C
  FILING DATE: April 4, 1994
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 07/868,354
  FILING DATE: April 10, 1992
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/745,206
  FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/620,250
  FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/482,384
  FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/603,751
  FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: WO PCT/US89/01408
  FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/176,899
  FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
  NAME: Seidman, Stephanie L.
  REGISTRATION NUMBER: 33,779
  REFERENCE/DOCKET NUMBER: 52516 (P519739)
  TELEPHONE: (619)238-0999
  TELEFAX: (619)238-0062
  INFORMATION FOR SEQ ID NO: 8:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 7175 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
    FEATURE:
      NAME/KEY: CDS
      LOCATION: 144..6857
    FEATURE:
      NAME/KEY: 5'UTR
      LOCATION: 1..143
    FEATURE:
      NAME/KEY: 3'UTR
      LOCATION: 6855..7175
US-08-223-305C-8

Query Match      3.8%; Score 49; DB 2; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0091;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GACGCTGCGTGCAGGTGAACCCGCGCGCGAGCGCTCGGTGTCAGTGTTCCTCGGGCGC 309
DB 2751 GAGGCCCCGAAAGGCGGAGCGGGGAGCGCCGCTGCGCGGAGGAGCGCGCGCGCGCAC 2810
QY 310 CGCAGCTGCAGCTGCAGGGTGCAGCGAGCCCGAGCGCCGATCGGTTCTGTGTCAA 369
DB 2811 CGCA-GCCACAGCAGAGGAGCGCGGGGCC--CCCGAGAGCGCGAGCGAGCGGGCGCA 2867
QY 370 CCCGTGTCACGCGCGCGCGGAGATCCCGCGATCTCGCAGACCGTAGCCCCCGTTC 429
DB 2868 GSCCCAGGCCCCGAGGCGCGCGCGGCGCACACCGCGCGGCTCCCGAGGAGGCGCGCC 2927
QY 430 TCGTCCGTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489
DB 2928 GAGCGGGAGCCCCGACGCGCACCGCGCGCACCGGACCGGATCCGAGAGGAGTGCGCC 2987

QY 490 CCCAGAGGAGGAGGAGCGCGCATCTCTGGGACCCCGGAACCCGAGCCGAGAGAG 549
DB 2988 GCGCCCAAGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 3047
QY 550 GTGGCCGCGAGGAAAGCGGTCTCCCGAGCGCGGAAAGCGAGGAGGCGGATGTTTCAGGCTGCA 609
DB 3048 GCGGAGAGCGGGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCT 3107
QY 610 GCGGAGGCGCGGTGGGAGCAGCAGCCACCGGAGGAGCCGGAACAGTGTGGCGGCGAT 668
DB 3108 CACGAGGCTGTGGAGAGGAGCACCAGGAGGAGGCGCGCGGAGAGGAGGAGGAGGCTGAGAT 3166

RESULT 11
US-08-149-097D-8
; Sequence 8, Application US/08149097D
; Patent No. 5874236
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/149,097D
; FILING DATE: 05-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/105,536
; FILING DATE: 11-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US92/06903
; FILING DATE: 14-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/914,231
; FILING DATE: 13-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/868,354
; FILING DATE: 10-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/620,250
; FILING DATE: 30-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/482,384
; FILING DATE: 20-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/603,751
; FILING DATE: 04-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US89/01408
; FILING DATE: 04-APR-1989
; APPLICATION NUMBER: US 07/176,899
```

```

; FILING DATE: 04-APR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-55038
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 238-0999
; TELEFAX: (619) 238-0062
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7175 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 144..6857
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..1143
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 6855..7175
;
; US-08-149-097D-8
;
; Query Match 3.8%; Score 49; DB 2; Length 7175;
; Best Local Similarity 48.0%; Pred.No. 0.0091;
; Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;
;
; QY 250 GAGCGTCGCGTGAGGTGAACCCGCGCGCGAGCGCTTCGGTGTGAGTGTTCACATCGGGCGC 309
; DB 2751 GAGGCCCGGAAGCGGAGAGCGGGAGCCCGGTGCCCCGGGAGGAGCGCGCGCGCGCAC 2810
;
; QY 310 CGCAGCGTGAGCGCTGCAGGGTGCAGAGCCGAGCCCGACGCCCGATCGGGTTCCTGTCAA 369
; DB 2811 CGCA-GCCACAGCAAGGAGGCGCGGGGCC--CCCGAGGGCGCGGAGCGGAGCGGCGCG 2867
;
; QY 370 CCCCGTGGCCAGCGCGCGCGCGGAGATCCCGCGCATCTCTGGCAGACCGTAGCCCCGTTCC 429
; DB 2868 GGCACAGCCCCGAGGGCGCGCGCGCACACCGCGCGGCTCCCCGGAGGAGCGGCC 2927
;
; QY 430 TCGTCCTGTGACCTTCTGTGGCTCTCTCTCACTGTGAGTTGCGGAGGAGGAGGAGGAGCA 489
; DB 2928 GAGCGGAGCGCCGACGCGCACCGCGCGCACCGGACACCGAGTCCGAGCAAGAGTGGCC 2987
;
; QY 490 CCCACGAGGAGAGGGGAGCGCGCGCATCTCTCGGGACCCCGGAAACCGGAGCCGAGAGAG 549
; DB 2988 GCGCCCAAGGGCGAGCGCGCGCGCGGCGCACCGCGGGGCCCCCGAGCGGGGCCCGGGAG 3047
;
; QY 550 GTGGCCCGAGGAAAGCGGTCCCCAGCCGGAACGAGGAGGGCGATGTTTCAGGCTGCA 609
; DB 3048 GCGGAGAGCGGGAGGAGCGCGCGCGGGCACCGGGCCCCGGCACAGGCGCAGCTGCT 3107
;
; QY 610 GGGCAGCGCGGTGGGAGCAGCACCCACCGGAGGACCGAAACAGTGTGGCGCGAT 668
; DB 3108 CACGAGGCTGTGGAGAGGAGACCAACCGAGAGGAGGCCACCGAGAGAGGCTGAGAT 3166
;
;
; RESULT 12
; US-08-949-386-8
; Sequence 8, Application US/08949386
; Patent No. 6090623
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: McCue, Ann
; APPLICANT: Gillespie, Alison
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
;

```


Qy 610 GGGCAGGCGGGTGGGAGCAGCAGCCACCCGAGGAGACCGGAACAGTGTGGCGCGCAT 668
Db 3108 CACGAGGCTGTGGAGAGGAGACCCACCGAGAGGAGGCCACCGAGAGGAGGCTGAGAT 3166

Search completed: January 5, 2005, 19:22:37
Job time : 137 secs

FILING DATE: 5-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/105,536
FILING DATE: 11-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/914,231
FILING DATE: 13-JULY-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: 10-APR-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/745,206
FILING DATE: 15-AUG-1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-519812
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 238-0999
TELEFAX: (619) 238-0062
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-450-272-8

Query Match 3.8%; Score 49; DB 3; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0091;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;
Qy 250 GACGTCGGTGCAGGTGAACCCGCGCGCAGCGCTCGGTGCAGTGTTCACCTCGGGCGC 309
Db 2751 GAGGCCCGAAGCGGAGAGCGGGGAGCCCGGTGCCCCGGAGAGCGGCCCGCGCGCAC 2810
Qy 310 CGCAGCGTGCAGCTGCAGGTGCGGAGCCAGCCCGCAGCCCGCATCGGGTTCCTGTCAA 369
Db 2811 CGCA-GCCACAGCAAGAGAGCGCGGGGCC--CCCGAGAGCGCGAGCGCGCGCGA 2867
Qy 370 CCCGTCGCCACGCGCGCGCGGAGATCCCGCGATCCTGGCAGACCGTAGCCCCGTTTC 429
Db 2868 GGCCCCAGGCCCGGAGGGCGCGCGGCGCACCCCGCGCGGCTCCCCCGAGAGGCGCGCC 2927
Qy 430 TCGTCCGTGACCTTCTGTGGCTCTCTCTCTCACTGGAGGTTGCGGAGGCGAGCGACACA 489
Db 2928 GAGCGGAGCCCCCGACGCCACCGCGCGCACCGGACACAGGATCCGAGCAAGGATGCGCC 2987
Qy 490 CCCACGAGGAGAGGGAGCCCGGCATCTTCGGGGACCCCGGGAACCGGAGCCGAGAGAG 549
Db 2988 GGCGCAAGGGCGAGCGCGCGCGCGCACCGCGCGGCCCGCGCGCGCGCGCGCGAG 3047
Qy 550 GTGGCGCGGAGGAAGCGGTCCCCCAGCGCGGAGAGAGGGGGGATGTTTCAGGCTGCA 609
Db 3048 CGGAGAGCGGGGAGGAGCGCGCGCGCGCACCGGGGCCCGGCAAGAGGCGCGCTGCT 3107

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2005, 18:07:17 ; Search time 745 Seconds
(without alignments)

9681.641 Million cell updates/sec

Title: US-09-830-810A-1

Perfect score: 1277

Sequence: 1 aaggcggcgagggcgcgga.....acaaaaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4176236 seqs, 2824127955 residues

Total number of hits satisfying chosen parameters: 8352472

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1277	100.0	1277	9	US-09-844-864-1
2	1234.2	96.6	1258	18	US-10-475-502-1
3	799.4	62.6	814	9	US-09-844-864-18
4	799.4	62.6	814	9	US-09-844-864-18
5	753.8	59.0	4090	18	US-10-475-502-11
6	752.2	58.9	809	9	US-09-844-864-22
7	752.2	58.9	809	9	US-09-844-864-22
8	260.6	20.4	375	9	US-09-844-864-25
9	134.6	10.5	2075	18	US-10-475-502-13
10	98.8	7.7	123	9	US-09-844-864-19
11	98.8	7.7	123	9	US-09-844-864-23
12	72	5.6	105	9	US-09-844-864-20

13	70.4	5.5	105	9	US-09-844-864-24
c 14	55.2	4.3	1041	17	US-10-437-963-8597
c 15	55	4.3	972	18	US-10-425-115-17653
16	54.2	4.2	1500	16	US-10-374-780A-680
17	54.2	4.2	1500	16	US-10-412-6998-1133
c 18	54.2	4.2	4667	18	US-10-723-860-5759
c 19	52.8	4.1	2664	17	US-10-437-963-52312
c 20	52.2	4.1	720	17	US-10-437-963-77205
c 21	51.6	4.0	987	16	US-10-084-846A-99
c 22	51.6	4.0	9816	16	US-10-084-846A-1
c 23	51.6	4.0	9816	16	US-10-084-846A-2
c 24	51	4.0	1185	15	US-10-156-761-6879
c 25	51	4.0	9025608	15	US-10-156-761-1
26	50.6	4.0	1190	17	US-10-437-963-36882
27	50.4	3.9	1388	17	US-10-767-701-14298
c 28	50.2	3.9	1673	16	US-10-374-780A-687
c 29	50.2	3.9	1673	16	US-10-412-6998-1140
30	49.6	3.9	400	15	US-10-193-002-174
31	49.6	3.9	400	15	US-10-084-843-179
32	49.6	3.9	14800	9	US-09-954-456-1601
33	49.6	3.9	14800	15	US-10-269-909-61
34	49.6	3.9	14800	17	US-10-717-597-183
35	49.6	3.9	14835	15	US-10-240-965-113
36	49	3.8	1092	17	US-10-767-701-9739
37	49	3.8	1839	11	US-09-758-759-150
38	49	3.8	5235	16	US-10-627-370-3
39	49	3.8	6792	16	US-10-627-370-1
40	49	3.8	7175	16	US-10-375-253-13
41	49	3.8	7177	13	US-10-033-026-7
42	49	3.8	7177	18	US-10-736-883-33
43	49	3.8	7362	16	US-10-375-253-11
44	49	3.8	7363	18	US-10-723-860-5192
45	49	3.8	7364	9	US-09-954-456-1179

ALIGNMENTS

RESULT 1

US-09-844-864-1
; Sequence 1, Application US/09844864
; Patent No. US20020042926A1
; GENERAL INFORMATION:
; APPLICANT: Matzuk, Martin
; APPLICANT: Ren, Xongsheng
; APPLICANT: Wu, Xuemei
; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
; FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
; CURRENT APPLICATION NUMBER: US/09/844,864
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/106,020
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: PCT/US99/25209
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1277
; TYPE: DNA
; ORGANISM: Mus musculus
; US-09-844-864-1

Query Match	100.0%	Score 1277;	DB 9;	Length 1277;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1277;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AAGCGGCGGAGGCGGCGGAGCGAGCCATGTTCCCGGAGCAGCGTTCCACCCCTGCCCG	60	
Db	1	AAGCGGCGGAGGCGGCGGAGCGAGCCATGTTCCCGGAGCAGCGTTCCACCCCTGCCCG	60	
Qy	61	CATCCTTATCCGAGGCGGAGCCCAAGCCGGGATGCTCGAGGCTTCGAGGCGGCGGCTGC	120	
Db	61	CATCCTTATCCGAGGCGGAGCCCAAGCCGGGATGCTCGAGGCTTCGAGGCGGCGGCTGC	120	

121 QY GCACCCGCGCCCTCTCTCTCCCGGCTACAGACAGCTCATGGCCGCGAGTACGTC 180
 121 Db CGACCCGCGCCCTCTCTCTCTCCCGGCTACAGACAGCTCATGGCCGCGAGTACGTC 180
 181 QY GACAGCCACAGCGGCGACAGCTCATGGCCGCTGCTGTCGCGATGGTCCCGGTCGGTC 240
 181 Db GACAGCCACAGCGGCGACAGCTCATGGCCGCTGCTGTCGCGATGGTCCCGGTCGGTC 240
 241 QY AGCAGCCGTGACCTGCGGTGAGTGAACCGCGCCGCGAGCTCGGTGAGTGTCA 300
 241 Db AGCAGCCGTGACCTGCGGTGAGTGAACCGCGCCGCGAGCTCGGTGAGTGTCA 300
 301 QY CTGCGGCGCGCAGCTGCGAGCTTCAGGGTGCAGGACGCCCGAGCTCGGGT 360
 301 Db CTGCGGCGCGCAGCTGCGAGCTTCAGGGTGCAGGACGCCCGAGCTCGGGT 360
 361 QY TCCTGTCAACCCGCTGGCCACCGCGCGCGGGAGATCCCGGATCTGGCAGACCGTA 420
 361 Db TCCTGTCAACCCGCTGGCCACCGCGCGCGGGAGATCCCGGATCTGGCAGACCGTA 420
 421 QY GCCCGCTTCTCGTCCGTGACCTTCTGTGGCTCTCTCTCTCACTGGAGTTGCGGAGGC 480
 421 Db GCCCGCTTCTCGTCCGTGACCTTCTGTGGCTCTCTCTCTCACTGGAGTTGCGGAGGC 480
 481 QY AGGCGACACCCACGAGGGAGGGAGCCCGGATCTCGGGACCGGGAACCGGAG 540
 481 Db AGGCGACACCCACGAGGGAGGGAGCCCGGATCTCGGGACCGGGAACCGGAG 540
 541 QY CCGAGAGAGTGGCCCGGAGGAAGCGTCCCGCAGCGAGGAGGAGGCGATGTT 600
 541 Db CCGAGAGAGTGGCCCGGAGGAAGCGTCCCGCAGCGAGGAGGAGGCGATGTT 600
 601 QY CAGGCTGCGAGGCGAGCCCGGTGGAGCAGCAGCCACCAACCGAGGACCGGAACAGTGTG 660
 601 Db CAGGCTGCGAGGCGAGCCCGGTGGAGCAGCAGCCACCAACCGAGGAGCGGAGTGTG 660
 661 QY GCGGCGATGAGTCTGAGCTGGAGCGAGGAGGAGTCTCTGCGCAGAGATGCTCAG 720
 661 Db GCGGCGATGAGTCTGAGCTGGAGCGAGGAGGAGTCTCTGCGCAGAGATGCTCAG 720
 721 QY GACCCGCTGATTCGATGCTCCCTCGAGACCAAGCCCTCCCGCAAGACCGGAGCAGGAC 780
 721 Db GACCCGCTGATTCGATGCTCCCTCGAGACCAAGCCCTCCCGCAAGACCGGAGCAGGAC 780
 781 QY AAGGAGCGCTCGGTTTCAGTTCTTAGAGCAGAGTACGGCTACTATCACTGCAAGGAC 840
 781 Db AAGGAGCGCTCGGTTTCAGTTCTTAGAGCAGAGTACGGCTACTATCACTGCAAGGAC 840
 841 QY TGCAAAATCCGTGGGAGCGCTATGTGTGTGTGTGTCAGGCGCACCAAGTAAAGTGTGA 900
 841 Db TGCAAAATCCGTGGGAGCGCTATGTGTGTGTGTGTCAGGCGCACCAAGTAAAGTGTGA 900
 901 QY CTTCAAAACAGTCTTCCGCGAGTGTGAGAAATCCTACAAACCTTTACAGAGTGAAGACAT 960
 901 Db CTTCAAAACAGTCTTCCGCGAGTGTGAGAAATCCTACAAACCTTTACAGAGTGAAGACAT 960
 961 QY CACTGTCAAAAGTGTAAAGAACTAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 961 Db CACTGTCAAAAGTGTAAAGAACTAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 1021 QY CCCTTAAAGCGCCCATCGCAAGACTGTGTGGAGATGCAAGGACAAACCGCTGTCTGCTG 1080
 1021 Db CCCTTAAAGCGCCCATCGCAAGACTGTGTGGAGATGCAAGGACAAACCGCTGTCTGCTG 1080
 1081 QY CGACAGCACTTACAGTCTCAAAATACATCATATTTAGTGAAGTCAAAACGTTTCTGTAGA 1140
 1081 Db CGACAGCACTTACAGTCTCAAAATACATCATATTTAGTGAAGTCAAAACGTTTCTGTAGA 1140
 1141 QY TGGGGCTAATGGATGACAGTGAAGTCTTCTCCCTCTTCACTCTTCTCTTCTTCTTCTTCTT 1200
 1141 Db TGGGGCTAATGGATGACAGTGAAGTCTTCTCCCTCTTCACTCTTCTCTTCTTCTTCTTCTT 1200

1201 QY TCTTCATGACAGACAGTGTACTTGGATATAAAGCCCTGTGAATAAAAGGTATTGCAACA 1260
 1201 Db TCTTCATGACAGACAGTGTACTTGGATATAAAGCCCTGTGAATAAAAGGTATTGCAACA 1260
 1261 QY AAAAAAAAAAAAAAAAAA 1277
 1261 Db AAAAAAAAAAAAAAAAAA 1277

RESULT 2
 US-10-475-502-1
 ; Sequence 1, Application US/104/5502
 ; Publication No. US20040254132A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Pei
 ; APPLICANT: Wu, Xuemei
 ; APPLICANT: Matzku, Martin M.
 ; APPLICANT: Bai, Yuchen
 ; TITLE OF INVENTION: Ovary-Specific Genes and Proteins
 ; FILE REFERENCE: P01925US7
 ; CURRENT APPLICATION NUMBER: US/10/475,502
 ; CURRENT FILING DATE: 2003-10-22
 ; PRIOR APPLICATION NUMBER: PCT US02/13245
 ; PRIOR FILING DATE: 2002-04-26
 ; PRIOR APPLICATION NUMBER: US 09/844,864
 ; PRIOR FILING DATE: 2001-04-27
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 1258
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; US-10-475-502-1

Query Match 96.6%; Score 1234.2; DB 18; Length 1258;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1250; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

3 QY GCGCGGCGAGCGCGGAGCGCACCCATGTTCCCGCGGAGCAGTTCACCCCTGCCGCA 62
 1 Db GCGCGGCGAGCGCGGAGCGCACCCATGTTCCCGCGGAGCAGTTCACCCCTGCCGCA 60
 63 QY TCCTTATCCGAGCGCACCAAGCGGCGGATGGCTGGAGTTCCGAGCCAGGGGCTCGG 122
 61 Db TCCTTATCCGAGCGCACCAAGCGGCGGATGGCTGGAGTTCCGAGCCAGGGGCTCGG 120
 123 QY ACCCGCGCCCTCTCTCTCCCGGCTACAGACAGCTCATGGCCGCGGAGTACGTGCA 182
 121 Db ACCCGCGCCCTCTCTCTCTCCCGGCTACAGACAGCTCATGGCCGCGGAGTACGTGCA 180
 183 QY CAGCCACAGCGGCGACAGCTCATGGCCCTCTCTGCGGATGGGTCCCGGTCGGTCAG 242
 181 Db CAGCCACAGCGGCGACAGCTCATGGCCCTCTCTGCGGATGGGTCCCGGTCGGTCAG 240
 243 QY CAGCGCTGACCTGCGGTGAGGTCAACCCCGCGGAGCGCTCGGTGAGTGTCACT 302
 241 Db CAGCGCTGACCTGCGGTGAGGTCAACCCCGCGGAGCGCTCGGTGAGTGTCACT 300
 303 QY CCGCGCGCGCGACGCTGACGCTGAGGGTGCAGGCGCGGAGCGGAGCGGCGGATCGGGTTC 362
 301 Db CCGCGCGCGCGACGCTGACGCTGAGGGTGCAGGCGCGGAGCGGAGCGGCGGATCGGGTTC 360
 363 QY CTGTCAACCCCGTGGCCACCGCGGCGCGGAGATCCCGGATCCCGGATCCCGGATCCCGGATCC 422
 361 Db CTGTCAACCCCGTGGCCACCGCGGCGCGGAGATCCCGGATCCCGGATCCCGGATCCCGGATCC 420
 423 QY CCGGTTCTGCTCGGTGAGCTTCTGTGGCTCTCTCTCTCTCACTGGAGTTGCGGAGGAGCAG 482
 421 Db CCGGTTCTGCTCGGTGAGCTTCTGTGGCTCTCTCTCTCTCACTGGAGTTGCGGAGGAGCAG 480
 483 QY GCAGACACCCACGAAAGGAGGAGGAGCCCGGATCTCTCGGAGACCCCGGAGCCCGGAGCC 542
 481 Db GCAGACACCCACGAAAGGAGGAGGAGCCCGGATCTCTCGGAGACCCCGGAGCCCGGAGCC 540


```
/ CURRENT APPLICATION NUMBER: US/10/475,502
/ CURRENT FILING DATE: 2003-10-22
/ PRIOR APPLICATION NUMBER: PCT US02/13245
/ PRIOR FILING DATE: 2002-04-26
/ PRIOR APPLICATION NUMBER: US 09/844,864
/ PRIOR FILING DATE: 2001-04-27
/ NUMBER OF SEQ ID NOS: 23
/ SOFTWARE: Patent in version 3.1
/ SEQ ID NO 11
/ LENGTH: 6873
/ TYPE: DNA
/ ORGANISM: mus musculus
US-10-475-502-11

Query Match      62.6%; Score 799.4; DB 18; Length 6873;
Best Local Similarity 99.9%; Pred. No. 7.4e-227;
Matches 800; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GCGGGCGAGGCGCGGAGCGACCCATGTTCCCGGCGAGCAGTTCACCCCTGCCGCA 62
DB 2513 GCGGGCGAGGCGCGGAGCGACCCATGTTCCCGGCGAGCAGTTCACCCCTGCCGCA 2572

QY 63 TCCTTATCCGACAGGCGACCAAAAGCCGGGATGGCTCGAGGTTCCGAGCCAGGGGCTGCCG 122
DB 2573 TCCTTATCCGACAGGCGACCAAAAGCCGGGATGGCTCGAGGTTCCGAGCCAGGGGCTGCCG 2632

QY 123 ACCCGCGCCCTCTCTTCTCCCGGCTACAGAGCTCATGGCGCGGAGTACGTGCA 182
DB 2633 ACCCGCGCCCTCTCTTCTCCCGGCTACAGAGCTCATGGCGCGGAGTACGTGCA 2692

QY 183 CAGCCACAGCGGCGACAGCTCATGSCCTCTGTCTCGCGATGGGTCCCGGTCCGTGAG 242
DB 2693 CAGCCACAGCGGCGACAGCTCATGSCCTCTGTCTCGCGATGGGTCCCGGTCCGTGAG 2752

QY 243 CAGCGGTGACGTTGCGGTGACGTAACCCCGCGCGACGCTCGGTGACGTGTTACT 302
DB 2753 CAGCGGTGACGTTGCGGTGACGTAACCCCGCGCGACGCTCGGTGACGTGTTACT 2812

QY 303 CGGGCGCGGACGCTGACGCTGAGGTTCCGAGCCAGCCCGGACCCCGATCGGGTTC 362
DB 2813 CGGGCGCGGACGCTGACGCTGAGGTTCCGAGCCAGCCCGGACCCCGATCGGGTTC 2872

QY 363 CTGTCAACCCGTTGGCCAGCGCGCGCGGAGATCCCGCGATCTTGGCAGACCGTAGC 422
DB 2873 CTGTCAACCCGTTGGCCAGCGCGCGGAGATCCCGCGATCTTGGCAGACCGTAGC 2932

QY 423 CCGGTTCTCGTTCGCTGACCTTCTGTGGCTCTCTCTCTCACTGGAGGTTGCGGGAGGCA 482
DB 2933 CCGGTTCTCGTTCGCTGACCTTCTGTGGCTCTCTCTCTCACTGGAGGTTGCGGGAGGCA 2992

QY 483 GCAGACACCCAGAGGAGGGAGCCCGGCATCTCGGGACCCCGGAAACCGGAGCC 542
DB 2993 GCAGACACCCAGAGGAGGGAGCCCGGCATCTCGGGACCCCGGAAACCGGAGCC 3052

QY 543 GAGAGAGGTGGCGCGGAGAAACCGGTTCGCCAGCCGCGAAGCAGGAGGGCGATTTCA 602
DB 3053 GAGAGAGGTGGCGCGGAGAAACCGGTTCGCCAGCCGCGAAGCAGGAGGGCGATTTCA 3112

QY 603 GGTGACAGGCGAGGCGGGTGGAGCAGCCGACACCCAGGAGGACCGGAACAGTGTGC 662
DB 3113 GGTGACAGGCGAGGCGGGTGGAGCAGCCGACACCCAGGAGGACCGGAACAGTGTGC 3172

QY 663 GCGGATGACGCTGAGCTGGGAGCGAGGAGCCGATGTCCTGCGCAGAGATGGCTCAGGA 722
DB 3173 GCGGATGACGCTGAGCTGGGAGCGAGGAGCCGATGTCCTGCGCAGAGATGGCTCAGGA 3232

QY 723 CCGCGGTGATTGGATGCCCCCTCGACACCAAGGCTCCCGCAAAAGCAGGAGCAGACAA 782
DB 3233 CCGCGGTGATTGGATGCCCCCTCGACACCAAGGCTCCCGCAAAAGCAGGAGCAGACAA 3292

QY 783 GGAGCGCTCGGTTCCAGTT 803
DB 3293 GGAGCGCTCGGTTCCAGGT 3313
```

```
RESULT 5
US-10-475-502-12
/ Sequence 12, Application US/10475502
/ Publication No. US20040254132A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Pei
/ APPLICANT: Wu, Xuemei
/ APPLICANT: Matzuk, Martin M.
/ APPLICANT: Bai, Yuchen
/ TITLE OF INVENTION: Ovary-Specific Genes and Proteins
/ FILE REFERENCE: P01925US7
/ CURRENT APPLICATION NUMBER: US/10/475,502
/ CURRENT FILING DATE: 2003-10-22
/ PRIOR APPLICATION NUMBER: PCT US02/13245
/ PRIOR FILING DATE: 2002-04-26
/ PRIOR APPLICATION NUMBER: US 09/844,864
/ PRIOR FILING DATE: 2001-04-27
/ NUMBER OF SEQ ID NOS: 23
/ SOFTWARE: Patent in version 3.1
/ SEQ ID NO 12
/ LENGTH: 4090
/ TYPE: DNA
/ ORGANISM: Mus musculus
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)..(4090)
/ OTHER INFORMATION: N equals unknown
US-10-475-502-12
```

```
Query Match      59.0%; Score 753.8; DB 18; Length 4090;
Best Local Similarity 97.5%; Pred. No. 2.3e-213;
Matches 781; Conservative 0; Mismatches 7; Indels 13; Gaps 1;

QY 3 GCGGGCGAGGCGCGGAGCGACCCATGTTCCCGGCGAGCAGTTCACCCCTGCCGCA 62
DB 1 GCGGGCGAGGCGCGGAGCGACCCATGTTCCCGGCGAGCAGTTCACCCCTGCCGCA 60

QY 63 TCCTTATCCGACAGGCGACCAAAAGCCGGGATGGGTTCGAGGTTCCGAGCCAGGGGCTGCCG 122
DB 61 TCCTTATCCGACAGGCGACCAAAAGCCGGGATGGGTTCGAGGTTCCGAGCCAGGGGCTGCCG 120

QY 123 ACCCGCGCCCTCTCTTCTCCCGGCTACAGACAGCTCATGCGCGCGGAGTACGTGCA 182
DB 121 ACCCGCGCCCTCTCTTCTCTCCCGGCTACAGACAGCTCATGCGCGCGGAGTACGTGCA 180

QY 183 CAGCCACAGCGGCGACAGCTCATGGCCCTCTGTCTCGCGATGGGTCCCGGTCCGTGAG 242
DB 181 CAGCCACAGCGGCGACAGCTCATGGCCCTCTGTCTCGCGATGGGTCCCGGTCCGTGAG 240

QY 243 CAGCCGTGACGCTGCGGTGCAAGTGAACCCCGCGCGAGCTCCGTGCAAGTGTTCAT 302
DB 241 CAGCCGTGACGCTGCGGTGCAAGTGAACCCCGCGCGAGCTCCGTGCAAGTGTTCAT 300

QY 303 CGGGCGCGGACGCTGACGCTGAGGTTCCGAGCAGCCCGGAGCAGCCCGATCGGGTTC 362
DB 301 CGGGCGCGGACGCTGACGCTGAGGTTCCGAGCAGCCCGGAGCAGCCCGATCGGGTTC 360

QY 363 CTGTCAACCCCGTGGCCAGCGCGCGCGGAGATCCCGCGATCTCTGGCAGACCGTAGC 422
DB 361 CTGTCAACCCCGTGGCCAGCGCGCGCGGAGATCCCGCGATCTCTGGCAGACCGTAGC 420

QY 423 CCGGTTCTCGTTCGCTGACCTTCTGTGGCTCTCTCTCTCACTGGAGGTTGCGGGAGGCA 482
DB 421 CCGGTTCTCGTTCGCTGACCTTCTGTGGCTCTCTCTCTCACTGGAGGTTGCGGGAGGCA 480

QY 483 GCAGACACCCAGAGGAGGGAGCCCGCATCTCTCGGGACCCCGGAAACCGGAGCC 542
DB 481 GCAGACACCCAGAGGAGGGAGCCCGCATCTCTCGGGACCCCGGAAACCGGAGCC 540

QY 543 GAGAGAGGTGGCGGAGGAAAGCGGTCCCGCAGCCGCGAAGCAGGAGGGCGATGTTCA 602
DB 541 GAGAGAGGTGGCGGAGGAAAGCGGTCCCGCAGCCGCGAAGCAGGAGGGCGATGTTCA 600
```

Db 541 GAGAGAGTGGCGGTGAGAAAGCGTCCCGAGCGCGAAGCGAGGAGGCGACGTTCA 600
Qy 603 GGCTGACGGAGGCGGGTGGAGCAGACGCCACCGGAGGACCGACAGTGTGC 662
Db 601 GGCTGACGGAGGCGGGTGGAGCAGACGCCACCGGAGGACCGACAGTGTGC 660
Qy 663 GGCGATGAGTCTGAGCCTGGAGCGAGGAGCCATGTCTCTCCGACAGATGGCTCAGGA 722
Db 661 GGCGATGAGTCTGAGCCTGGAGCGAGGAGCCATGTCTCTCCGACAGATGGCTCAGGA 720
Qy 723 CCCCCTGATTCGGATGTCCTCGAGACAGGCTCTCCCGCAAGACGAGGAGGACAA 782
Db 721 CCCCCTGATTCGGATGTC-----CCTCCCGCAAGACCAAGCAGGACAA 767
Qy 783 GGAGCGCTGCGTTTCCAGTT 803
Db 768 GGAGCTCTGCTTCCAGGT 788

RESULT 6

US-09-844-864-22
; Sequence 22, Application US/09844864
; Patent No. US20020042926A1
; GENERAL INFORMATION:
; APPLICANT: Matzuk, Martin
; APPLICANT: Ren, Yongsheng
; APPLICANT: Wu, Xuemei
; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
; FILE REFERENCE: P0192502 / 09807797 / OTA 99-48
; CURRENT APPLICATION NUMBER: US/09/844,864
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/106,020
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: PCT/US99/25209
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 809
; TYPE: DNA
; ORGANISM: mus musculus
US-09-844-864-22

Query Match 58.9%; Score 752.2; DB 9; Length 809;
Best Local Similarity 97.1%; Pred No. 3.4e-213;
Matches 778; Conservative 0; Mismatches 19; Indels 5; Gaps 1;
Qy 3 GGCGGGGAGGCGGGGACGACCCATGTTCCCGGAGCAGCTTCCACCCCTGCCGCA 62
Db 1 GGCGGGGAGGCGGGGACGACCCATGTTCCCGGAGCAGCTTCCACCCCTGCCGCA 60
Qy 63 TCCTTATCCGAGGCGACCAAGCGGGATGGCTGGAGTTCCGAGCAGGCGTCCG 122
Db 61 TCCTTATCCGAGGCGACCAAGCGGGATGGCTGGAGTTCCGAGCAGGCGTCCG 120
Qy 123 ACCGCGCCCTCTCTTCTCCCGGTACAGACAGCTCATGGCCGGAGTACGTCGA 182
Db 121 ACCGCGCCCTCTCTTCTCCCGGTACAGACAGCTCATGGCCGGAGTACGTCGA 180
Qy 183 CAGCCACAGCGGGACAGCTCATGGCCCTGCTGTCGGATGGGTCCCGGTCCGTCAG 242
Db 181 CAGCCACAGCGGGACAGCTCATGGCCCTGCTGTCGGATGGGTCCCGGTCCGTCAG 240
Qy 243 CAGCGGTACCGTGGGTGAGGTGAACCCCGCGCGACGCTCGGTGAGTGTCACT 302
Db 241 CAGCGGTACCGTGGGTGAGGTGAACCCCGCGCGACGCTCGGTGAGTGTCACT 300
Qy 303 CGGCGCGCAGCTGACGCTGACGGGTGCGAGCGAGCCCGACGCGCGGTTC 362
Db 301 CGGCGCGCAGCTGACGCTGACGGGTGCGAGCGAGCCCGACGCGCGGTTC 360
Qy 363 CTGTCAACCCCGTGGCAGCGCGCGCGGAGATCCCGCGATCTCTGGCAGACCGTAGC 422

Db 361 CTGTCAACCCCGTGGCAGCGCGCGCGGAGATCCCGCGATCTCTGGCAGACCGTAGC 420
Qy 423 CCGCTTCTCGTCCGTGACCTTCTGTGGCTTCTCTCTCTCACTGGAGGTTGCGGAGGCGAG 482
Db 421 CCGCTTCTCGTCCGTGACCTTCTGTGGCTTCTCTCTCTCACTGGAGGTTGCGGAGGCGAG 480
Qy 483 GCAGACACCCACGAAAGGAGGAGGAGCCCGGCAATCTCTGGGGACCCCGGAAACCGGAGCC 542
Db 481 GCAGACACCCACGAAAGGAGGAGGAGCCCGGCAATCTCTGGGGACCCCGGAAACCGGAGCC 540
Qy 543 GAGAGAGTGTCGCGGAGGAAAGCGGTCCCGACCGCGAAGCGAGGAGGCGATGTTCA 602
Db 541 GAGAGAGTGTCGCGGAGGAAAGCGGTCCCGACCGCGAAGCGAGGAGGCGACGTTCA 600
Qy 603 GGCTGACGGGAGGCGGGTGGAGCAGCAGCACCACCGAGGACCGACACAGTGTGC 662
Db 601 GGCTGACGGGAGGCGGGTGGAGCAGCAGCACCACCGAGGACCGGAAACAGTGTGC 660
Qy 663 GGCGATGAGTCTGAGCCTGGAGCGAGGAGCCATGTCTCTCCGACAGATGGCTCAGGA 722
Db 661 GGCGATGAGTCTGAGCCTGGAGCGAGGAGCCATGTCTCTCCGACAGATGGCTCAGGA 720
Qy 723 CCCCCTGATTCGGATGTCCTCGAGACGAGGCTCTCCCGCAAGACGAGGAGGACAA 782
Db 721 CCCCCTGATTCGGATGTCCT-----CCCGCAAGCACCATACCTCTGACGAGGACAA 775
Qy 783 GGAGCGCTGCGTTTCCAGTT 803
Db 776 GGAGCTCTGCTTCCAGGT 796

RESULT 7

US-09-844-864-21
; Sequence 21, Application US/09844864
; Patent No. US20020042926A1
; GENERAL INFORMATION:
; APPLICANT: Matzuk, Martin
; APPLICANT: Ren, Yongsheng
; APPLICANT: Wu, Xuemei
; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
; FILE REFERENCE: P0192502 / 09807797 / OTA 99-48
; CURRENT APPLICATION NUMBER: US/09/844,864
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/106,020
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: PCT/US99/25209
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 305
; TYPE: DNA
; ORGANISM: mus musculus
US-09-844-864-21

Query Match 21.4%; Score 273.6; DB 9; Length 305;
Best Local Similarity 95.3%; Pred No. 7.8e-71;
Matches 282; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Qy 965 TGTCAAGTTGTAAAGAACTAGATGTGCTGCCAGTCAGATTCGCCACGTCGACCT 1024
Db 10 TTTCAAGTTGTAAAGAACTAGATGTGCTGCCAGTCAGATTCGCCACGTCGACCT 69
Qy 1025 AAAGCCCGCCATCGGCAAGACTTGTGGGAGATGCAAGGCAAAACGCTCTCTGGAC 1084
Db 70 AAAGCCCGCCATCGGCAAGACTTGTGGGAGATGCAAGGCAAAATGCTTCTCTGGAC 129
Qy 1085 AGACCTTTCAGCTTCAATATCATCTTTAGTGAGATCGAAACGTTTCTCTAGATGGG 1144
Db 130 AGACCTTTCAGCTTCAATATCATCTTTAGTGAGATCGAAACGTTTCTCTAGATGGG 189
Qy 1145 GCTAATGGAATGGCAAGAGGCTTCTCCCTCTTCCCTCTTCCCTTTCCAAATCTT 1204

Db 190 GCTATGATGACGACGAGCTTCTCCCTCTTCCCTCTTCCATTTCCAAATCTT 249
 Qy 1205 CATGACAGACGTGTACTTGGATATAAAGCTGTGAATAAAGGTATTGCAAAACA 1260
 Db 250 CATGACAGACGTGTACTTGGATATAAAGCTGTGAATAAAGGTATTGCAAAACA 305

RESULT 8

US-09-844-864-25
 ; Sequence 25, Application US/09844864
 ; Patent No. US20020042926A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Matzuk, Martin
 ; APPLICANT: Ren, Yongsheng
 ; APPLICANT: Wu, Xuemei
 ; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
 ; FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
 ; CURRENT APPLICATION NUMBER: US/09/844,864
 ; CURRENT FILING DATE: 2001-04-27
 ; PRIOR APPLICATION NUMBER: 60/106,020
 ; PRIOR FILING DATE: 1998-10-28
 ; PRIOR APPLICATION NUMBER: PCT/US99/25209
 ; PRIOR FILING DATE: 1999-10-28
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 25
 ; LENGTH: 375
 ; TYPE: DNA
 ; ORGANISM: mus musculus
 US-09-844-864-25

Query Match 20.4%; Score 260.6; DB 9; Length 375;
 Best Local Similarity 96.7%; Pred. No. 6.3e-67;
 Matches 266; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 Qy 986 AGATGTGCTGCCAGTCAGATTTCGCCACGTGGACCTTAAACGCCGCCCATCGCAAGAC 1045
 Db 101 AGATGTGCTGCCAGTCAGACTCGCCACGTGTACCTTAGACGCCGCCCATCGCAAGAC 160
 Qy 1046 TTGTGTGGAGATGCAAGACAAACGCTGTCTCGGACAGACACCTTCAGCTTCAAATAC 1105
 Db 161 TTGTGTGGAGATGCAAGACAAACGCTGTCTCGGACAGACACCTTCAGCTTCAAATAC 220
 Qy 1106 ATCATTTAGTGAGTCGAAACGTTTCTGCTAGATGGGGCTTAATGGATGCAAGTGA 1165
 Db 221 ATGATTTAGTGAGTCGAAACGTTTCTGCTAGATGGGGCTTAATGGATGCAAGTGA 280
 Qy 1166 GCTTTCTCCCTCTTCCACCTCTTCCCTTTCCAAATTTCTTCATGACAGACAGTGTACTTG 1225
 Db 281 GCTTTCTCCCTCTTCCACCTCTTCCCTTTCCAAATTTCTTCATGACAGACAGTGTACTTG 340
 Qy 1226 GATATAAGCCTGTGAATAAAGGTATTGCAAAACA 1260
 Db 341 GATATAAGCCTGTGAATAAAGGTATTGCAAAACA 375

RESULT 9

US-10-475-502-13
 ; Sequence 13, Application US/10475502
 ; Publication No. US20040254132A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Pei
 ; APPLICANT: Wu, Xuemei
 ; APPLICANT: Matzuk, Martin M.
 ; APPLICANT: Bai, Yuchen
 ; TITLE OF INVENTION: Ovary-Specific Genes and Proteins
 ; FILE REFERENCE: P01925US7
 ; CURRENT APPLICATION NUMBER: US/10/475,502
 ; CURRENT FILING DATE: 2003-10-22
 ; PRIOR APPLICATION NUMBER: PCT US02/13245
 ; PRIOR FILING DATE: 2002-04-26
 ; PRIOR APPLICATION NUMBER: US 09/844,864
 ; PRIOR FILING DATE: 2001-04-27

NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 13
 ; LENGTH: 2075
 ; TYPE: DNA
 ; ORGANISM: human
 US-10-475-502-13

Query Match 10.5%; Score 134.6; DB 18; Length 2075;
 Best Local Similarity 71.1%; Pred. No. 4.4e-29;
 Matches 207; Conservative 0; Mismatches 79; Indels 5; Gaps 2;
 Qy 968 CAAAGTTGTAACAACTAGATGTGCTGCCAGTCAGATTTGCCACGTCGACCCCTAAA 1027
 Db 1360 CAGAGTTGTAACAAACGAGATGTTCTGCCAGTAAACTTCCACGTCGACCCCTAAA 1419
 Qy 1028 CGCCCCCATCGGAAGACTTGTGTGGAGATGCAAGACAAACGCTGTCTCGGACAGC 1087
 Db 1420 CGGCCCCACCGTCAAGATTGTGCGGTAGATGCAAGGCAACGCTGTCTGTGACAGC 1479
 Qy 1088 ACCTTCAGCTTCAATATCATCTTTA-GTGAGAGTGCAGAAACGTTTCTGTAGATGGGC 1146
 Db 1480 ACTTTCAGCTTCAATATCATCTTTAGTGAAGTCACTGTGTGTGTCATGCCCTGATG 1539
 Qy 1147 TAATGGAATGACAAAGTCAGCTTTCTCCCTCTTCCCTCTTCCCTTTCCAAATTTCTCA 1206
 Db 1540 GAGTAGACGAGTCAGCTTTTCCGTGCTCTCTCCACCTCTCCCTTCTCAAAATACTTCA 1599
 Qy 1207 TGACAGACAGTGTACTTGGATATAAAGCCTGTGTAATAAAGGTATTGCAA 1257
 Db 1600 TGAAGGCAAGTGTATTCTG----AAAAAGCCTTCAAAATAAAGGTATTGCAA 1646

RESULT 10

US-09-844-864-19
 ; Sequence 19, Application US/09844864
 ; Patent No. US20020042926A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Matzuk, Martin
 ; APPLICANT: Ren, Yongsheng
 ; APPLICANT: Wu, Xuemei
 ; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
 ; FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
 ; CURRENT APPLICATION NUMBER: US/09/844,864
 ; CURRENT FILING DATE: 2001-04-27
 ; PRIOR APPLICATION NUMBER: 60/106,020
 ; PRIOR FILING DATE: 1998-10-28
 ; PRIOR APPLICATION NUMBER: PCT/US99/25209
 ; PRIOR FILING DATE: 1999-10-28
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 19
 ; LENGTH: 123
 ; TYPE: DNA
 ; ORGANISM: mus musculus
 US-09-844-864-19

Query Match 7.7%; Score 98.8; DB 9; Length 123;
 Best Local Similarity 98.0%; Pred. No. 6.1e-19;
 Matches 100; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 795 TTTCAGTTCTTAGACGAGAAAGTACGGCTACTATCACTGCAAGGACTGCAAAATCCGGTG 854
 Db 9 TGTTCAGTTCTTAGACGAGAAAGTACGGCTACTATCACTGCAAGGACTGCAAAATCCGGTG 68
 Qy 855 GGAGAGCGCCTATGT 896
 Db 69 GGAGAGCGCCTATGT 110

RESULT 11

US-09-844-864-23
 ; Sequence 23, Application US/09844864


```
Qy 209 CCTGCTCTCGGGATGGGTCCCGGTGCGTTCAGACGGTGAAGCTGCGTGCAGGTGA 268
Db 464 GCGCGGGCGGGCGGGTGGCGGCCCGGGCCCGGCCCGCGCGCGCGCGCGCGCG 405
Qy 269 ACCCGCGCGCGACG-----CCTCGGTGCGTGTTCAGTCCGGCGCGCGCGCGCG 322
Db 404 CCGCGGGTGGCGCGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 345
Qy 323 CTGCAAGGGTGGCGAGCGCCGACGCGCCGATCGGGTTCCTGTCAAACCCCGTGGCCACG 382
Db 344 GGGCGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 285
Qy 383 CCGCGCGCGGGAGATCCCGCGGATCCTCGGAGACCGTAGACCGCGTTCGTGCGTGA 442
Db 284 GGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 225
Qy 443 TCTGTGGCTCTCTCTCTCACTGGAGGTTGCGGGAGGCGAGACACCGAGGGAG 502
Db 224 GGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 165
Qy 503 AGGGGAGCGCGCGCATCTCTCGGGACCCGGGAACCGGAGCCGAGAGAGTGGCGCGAGGA 562
Db 164 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 105
Qy 563 AAGCGGTCCCGCAGCGCGAGCGAGGAGGCGATGTTGAGGCTGCGAGGCGCGCGCGGT 622
Db 104 CCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 45
Qy 623 GGGAGCAGCGCACACAC 640
Db 44 GCGGGCGCGCGCGCGCG 27
```

```
RESULT 15
US-10-425-115-17653/c
; Sequence 17653, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 17653
; LENGTH: 972
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(972)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_116102C.1
US-10-425-115-17653
```

```
Query Match 4.3%; Score 55; DB 18; Length 972;
Best Local Similarity 46.6%; Pred. No. 1.6e-05;
Matches 269; Conservative 0; Mismatches 304; Indels 4; Gaps 3;

Qy 21 CGCACCATTGTTCCCGCGGAGCAGTTCACACCTCGCGCGATCCTTATCCGAGGCCAC 80
Db 819 CACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 760
Qy 81 CAAAGCGGGATGGTGGAGGTTGCGAGCGAGGGGCTGCCGACCGCGCGCGCGCGCGCTT 140
Db 759 CACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 700
```

```
Qy 141 CCTCCCGCGCTACAGACAGCTCATGCGCGGAGTACGTGCG--ACAGCCACACGAGCGGCA 198
Db 699 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 640
Qy 199 CAGTCTATGCGCCTGCTGTCGCGGATGGGTCCCGGTGCGTTCAGCAGCGCGTGCAG 258
Db 639 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 580
Qy 259 GTGAGGTGAAC--CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 317
Db 579 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 520
Qy 318 GCAGCTGACGGGTGCGGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGG 377
Db 519 GGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 460
Qy 378 CCAAGCGCGCGCGCGGAGATCCCGCGGATCCTGCGAGACCGTAGCGCGCGTTCGTCCGT 437
Db 459 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 400
Qy 438 GACCTTCTGTGGCTCTCTCTCACTGGAGGTTGCGGGAGGCGAGCAGACACCGAA 497
Db 399 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 340
Qy 498 GGGAGAGGGGAGCGCGCGCATCTCTCGGGGACCCCGGGAACCGGAGCGGAGAGTGGCGCG 557
Db 339 GGCAG--GCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 281
Qy 558 GAGGAAAGCGGTTCCTCCCGCGCGCGCGCGAGCGAGGAGGGCG 594
Db 280 GGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 244
```

Search completed: January 5, 2005, 21:16:57
Job time : 756 secs